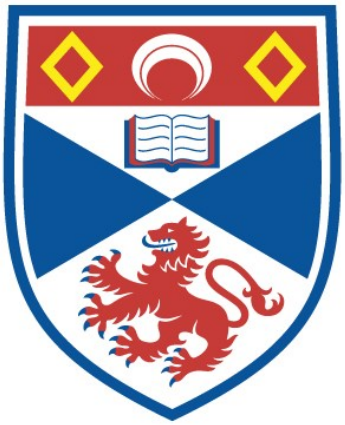
MULTIDIMENSIONAL AND HETERONUCLEAR NMR INVESTIGATIONS OF OLIGOSACCHARIDES IN THE FREE AND PROTEIN BOUND STATES

Mark John Milton

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1997

Multidimensional and Heteronuclear NMR Investigations of Oligosaccharides in the Free and Protein Bound States.

Mark John Milton

A thesis submitted for the degree of Doctor of Philosophy.

Centre for Biomolecular Sciences
University of St. Andrews

June 1997.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Table of Contents

Acknowledgments

I would like to thank my supervisor Steve Homans for all his help and encouragement throughout my Ph. D. project.

I would also like to thank Julia Richardson, Trevor Rutherford, Charles Weller and Richard Harris for their advice and assistance particularly with the NMR experiments. The synthesis of the carbon-13 enriched sugars was only possible with the support of Mark Probert in the laboratory, and Martek Biosciences for the gifts of fully labelled starting materials. In addition, I would like to acknowledge Bernard Badet for his kind donation of the glucosamine 6-phosphate synthase enzyme.

Finally I have to thank Douglas Low for keeping me entertained in the laboratory and Deepak Sharma, David Budd, and Gumminder Bassi for their continued support.

This work was funded by the BBSRC.

Declaration

I, Mark John Milton, hereby certify that this thesis, which is approximately 45 000 words in length, has been written by me, that it is a record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

17th June 1997 Mark John Milton

I was admitted as a research student in October 1993 and as a candidate for the degree of Doctor of Philosophy in October 1993; the higher study for which was carried out in the University of St. Andrews between October 1993 and August 1997.

17th June 1997 Mark John Milton

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree.

17th June 1997 Steve William Homans

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be supplied to any *bona fide* library or research worker.

17th June 1997 Mark John Milton

Abstract

The three dimensional structure and dynamics of three biologically important oligosaccharides were probed using a range of NMR techniques. The solution structure of oligosaccharide moiety of G_{M1} was solved using conventional ^1H - ^1H ROESY data in conjunction with molecular dynamics simulations. It was evident from this study that the dominant conformation in free solution is similar to the structure of oligo-G_{M1} in complex with the B-subunit of the enteropathogenic cholera toxin.

The solution dynamics of an analogue of globoside (Gb_4-OEtTMS), a cellular receptor for an *E. coli* verotoxin, was also elucidated using conventional ^1H - ^1H ROESY data. However, it was not clear that these data alone could adequately model the structure. Additional distance restraints were derived from ^1H - ^1H NOE contacts between exchangeable and non-exchangeable protons measured in H_2O/acetone-d_6 solutions at low temperature (-15 °C). It was concluded on the basis of these studies that Gb_4-OEtTMS is an extremely flexible sugar, however, the time-averaged conformation was similar to the proposed conformation of Gb_4 in complex with a verotoxin (VT2e).

In order to increase the structural information which may be measured from NMR spectroscopy, carbon-13 enriched carbohydrate moieties, commonly found in glycolipids and glycoproteins, were synthesised. Isotopically labelled *N*-acetyl neuraminic acid, α2,3 sialyl *N*-acetyl lactosamine (required for the synthesis of sialyl Lewis

^{x} antigen) and α2,3 sialyl lactose (oligo-G_{M3}), the core ganglioside which is required for the synthesis of more complex gangliosides (e.g. G_{M1}), was accomplished using a mixture of enzymatic and synthetic methods.

To test the applicability of heteronuclear NMR spectroscopy in the study of oligosaccharides, the solution structure and dynamics of [U-^{13}C] oligo-G_{M3} was probed by recording three dimensional ^{13}C-edited ROESY spectra, and long-range carbon-carbon and carbon-proton spin-coupling constants. Inter-residue ROEs could be unambiguously assigned from the 3D experiment, and these assignments contradict previous inter-residue ROE contacts published in the literature (based only on homonuclear data). Two additional inter-residue ROE contacts could be measured across the Galβ1-4Glc linkage. The solution conformation of [U-^{13}C] oligo-G_{M3} in complex with wheat germ agglutinin was solved on the basis of 3D TRNOESY-HSQC data. Theoretical TRNOE intensities back-calculated from full relaxation matrix calculations performed on the X-ray crystal structures of the complex were consistent with experimentally measured values and confirmed the flexible nature of the Galβ1-4Glc linkage when bound to the protein.

Finally, carbon-13 enriched *N*-acetyl neuraminic acid and galactose were introduced into the glycan chains of an intact glycoprotein FcREA. ^{13}C-^{1}H HSQC data suggest that the two arms of the biantennary glycan experience two distinct magnetic environments.

Table of Figures

Chapter 1.

Chapter 2.

Chapter 3.

List of Tables

Chapter 1.

Chapter 2.

Chapter 3.

Chapter 4.

Chapter 5.

Chapter 6.

COSY	Correlated Spectroscopy.
fs	femtosecond.
HMQC	Heteronuclear Multiple Quantum Coherence Spectroscopy.
HOHAHA	Homonuclear-Hartmann-Hahn.
HSQC	Heteronuclear Single Quantum Coherence Spectroscopy.
LRCC	Long Range Carbon-Carbon Coupling Spectroscopy.
NMR	Nuclear Magnetic Resonance.
NOE	Nuclear Overhauser Effect.
NOESY	Nuclear Overhauser Effect Spectroscopy.
ns	nanosecond.
OEtTMS	*O*-ethyltrimethylsilane.
ROE	Rotating Frame Nuclear Overhauser Effect.
ROESY	Rotating Frame Nuclear Overhauser Effect Spectroscopy.
ps	picosecond.
p-NP	*para*-nitrophenol.
TOCSY	Total Correlated Spectroscopy.
TRNOE	Transferred Nuclear Overhauser Effect.
TRNOESY	Transferred Nuclear Overhauser Effect Spectroscopy.
TRROESY	Transferred Rotating Frame Nuclear Overhauser Effect Spectroscopy.
TSP	3-(trimethylsilyl) propionic-*2,2,3,3,-d4*-acid, sodium salt.
[U-^{13}C]	Uniformly carbon-13 enriched (excluding the *N*-acetyl group of acetylated sugars).
WGA	Wheat Germ Agglutinin.

Symbols

^1H	Hydrogen atom (proton).
^{13}C	Carbon-13 atom.
$^nJ_{CH}$	n-bond carbon-proton coupling.
$^nJ_{CC}$	n-bond carbon-carbon coupling.
$^nJ_{HH}$	n bond proton-proton coupling.
B_0	External applied magnetic field.
δ	Chemical shift.
ϕ	Dihedral angle H1-C1-O1-Cx in a glycosidic linkage.
ϕ_1	Dihedral angle H6-C6-C7-H7 in a glycerol sidechain.
ϕ_2	Dihedral angle H7-C7-C8-H8 in a glycerol sidechain.
ϕ_3	Dihedral angle H8-C8-C9-O9 in a glycerol sidechain.
γ	Magnetogyric ratio.
$J(\omega)$	Spectral density.
k	Reduced rate constant.
k_{off}	Exchange off-rate.
k_{on}	Exchange on-rate.
K_D	Equilibrium dissociation constant.
ψ	Dihedral angle C1-O1-Cx-Hx in a glycosidic linkage.
τ_c	The correlation time for molecular reorientation.
τ_m	NOESY/ROESY mixing time.
$\tau_{m'}$	Residence time.
t_1	Acquisition time (1D NMR experiment); first evolution period (*N* D NMR experiment).

t_2	Acquisition time (2D NMR experiment); second evolution period (N D NMR experiment).
T_1	Time constant of spin-lattice (longitudinal) relaxation
T_2	Time constant of spin-spin (transverse) relaxation.

Chapter 1

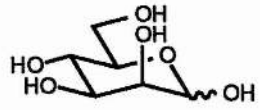
Introduction

1.1 Biological Function of Oligosaccharides

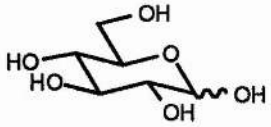
It is know well established that carbohydrate recognition plays an essential role in biological systems (reviewed in Varki, 1993). However, until the late 1960's the biological function of carbohydrates was generally thought to be limited to energy production and storage, and as a structural material. The possibility that carbohydrates may play a role in cellular recognition and differentiation was first suggested by Landsteiner's discovery of the ABH blood group antigens. Burnet provided the first evidence that oligosaccharides can function as cellular receptors for proteins when he demonstrated that a carbohydrate moiety (*N*-acetyl neuraminic acid, *See Figure 1.1*) on the surface of erythrocyte cells was essential for the attachment and entry of the influenza virus. Research over the last 30 years has know expanded the range of biological processes in which carbohydrates have been shown to play a pivotal role, for example, cellular adhesion, fertilisation, hormone action, cytolysis, metastasis, cell-cell recognition, toxin adhesion and invasion, immune evasion and activation, leucocyte extravasation, and the clearance of glycoproteins from the circulation by hepatic or reticuloendotheial cells.

Oligosaccharides involved in cellular recognition are normally found linked to a ceramide tail (i.e., glycolipids) or covalently attached to the side chains of specific amino acids such as asparagine (Asn), serine (Ser), and threonine (Thr) (i.e., glycoproteins). The carbohydrate moieties of these glycoconjugates may range in size from 2 to 20 residues and normally consist of neutral monosaccharides such as D-mannose (Man), D-glucose (Glc) and D-galactose (Gal), amino sugars such as *N*-acetyl-D-galactosamine (GalNAc) and *N*-acetyl-D-glucosamine (GlcNAc), the C-6-

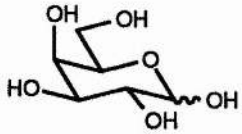
deoxy sugar L-fucose (Fuc), and the nine-carbon-atom acidic sugars of the sialic acid family of which *N*-acetyl neuraminic acid (Neu5Ac) is the most commonly found representative (Schauer, 1982)(for structures, *See Figure 1.1*).



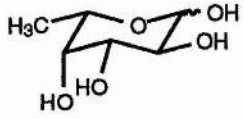
Mannose (Man)



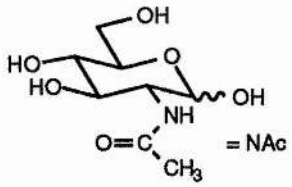
Glucose (Glc)



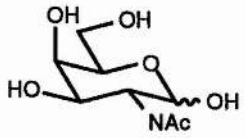
Galactose (Gal)



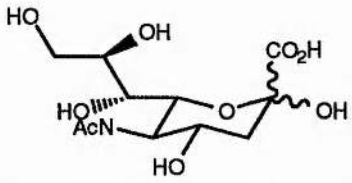
Fucose (Fuc)



N-acetyl glucosamine (GlcNAc)



N-acetyl galactosamine (GalNAc)



N-acetyl neuraminic acid (Neu5Ac)
The most common member of the sialic acid (Sia) family

Figure 1.1: Schematic representation of the typical constituent monosaccharide units of glycoproteins and glycosphingolipids.

In contrast to nucleotides in nucleic acids and amino acids in proteins, which can only interconnect in one way, the monosaccharide units in oligosaccharides can bond together in multiple ways to form a vast repertoire of compounds using only a few simple sugars. It is this potential for structural diversity that has made sugars a carrier of biological information i.e., coding for the health and type of a cell, and it's

organisation within tissue. This biological information is deciphered by a class of proteins referred to as lectins. Lectins normally exhibit exquisite specificity for their target oligosaccharide although the biological function of many (particularly plant lectins) is unknown at present. Other classes of proteins that interact with sugars include enzymes (involved in metabolism, catabolism, and sugar-nucleotide, mono and oligosaccharide synthesis), antibodies, permase proteins (involved in sugar transport), and DNA and RNA binding proteins.

Despite the potential complexity of oligosaccharide chains, glycosphingolipids and the *N*-linked glycoprotein glycans commonly found in animals can be classified into a few distinct groups. Glycolipids may be arranged into the following sets (*See Figure 1.2*): "ganglio"-, "globo"-, "isoglobo"-, "lacto"-, and "neo-lacto"- series.

Ganglio	
Galβ1-3GalNAcβ1-(Neu5Acα2-3)-4Galβ1-4GlcCer	**G_{M1}**
Globo	
GalNAcβ1-3Galα1-4Galβ1-4GlcCer	**globoside, Gb_4**
Isoglobo	
GalNAcβ1-3Galα1-3Galβ1-4GlcCer	**isogloboside**
Lacto	
Galβ1-3GlcNAcβ1-3Galβ1-4GlcCer	**lacto-*N*-tetraose**
Neo-lacto	
Galβ1-4GlcNAcβ1-4Galβ1-4GlcCer	**lacto-*N*-neo tetraose**

Figure 1.2: Examples of the core structures of animal glycosphingolipids.

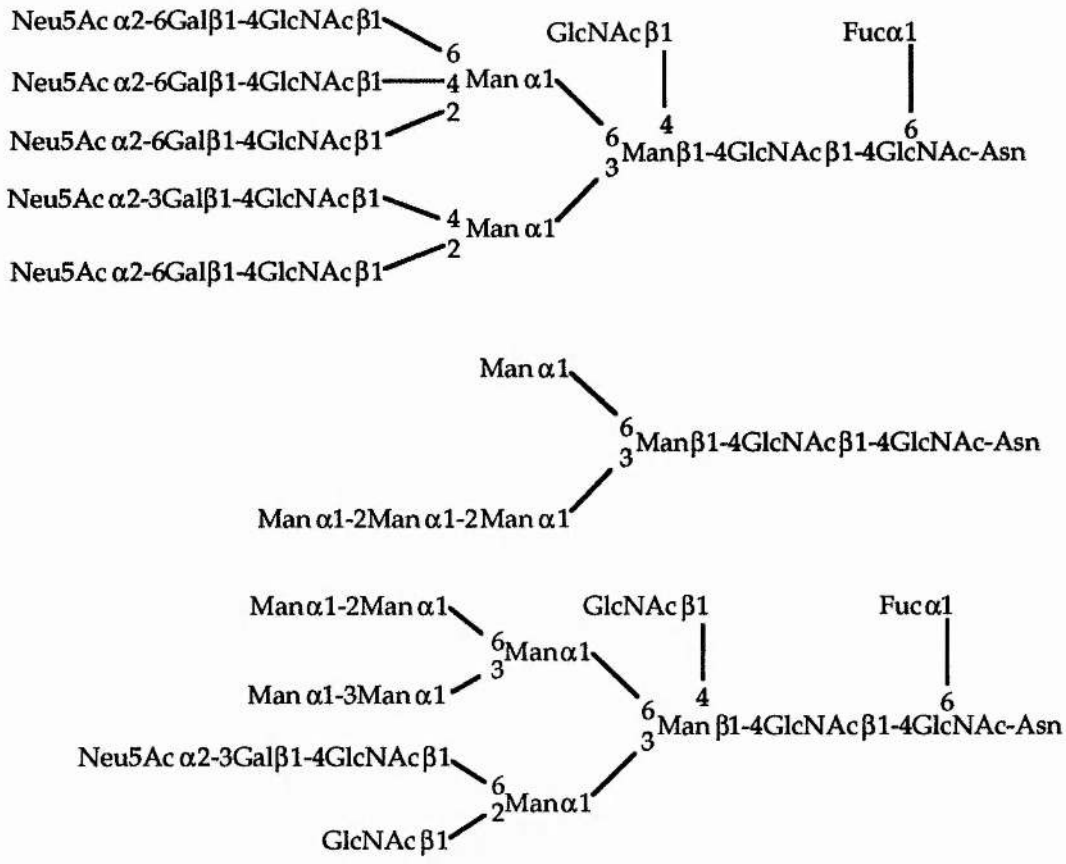


Figure 1.3: Examples of the three major subgroups of N-linked sugar chains: (1) complex type; (2) high mannose type; and (3) hybrid type.

The ganglio series glycolipids are differentiated from the other groups by the presence of at least one Neu5Ac residue per molecule. *N*-acetyl neuraminic acid is a uniquely functional sugar and is arguably the most important component of many glycoconjugates. Sialyloligosaccharides serve as cell surface determinants for influenza and other viruses, for mycoplasma, for blood group and tumour specific antibodies, for interferon, for recruiting lymphocytes seeking capillary sites for entry to the lymph system, for bacterial toxins, for cellular invasion and immune evasion by parasites, and for a variety of plant and animal lectins

(Reviewed in (Schauer, 1982; Reuter and Gabius, 1996; Varki, 1997). *N*-linked (asparagine-linked) glycoprotein glycans can generally be fitted into three broad families: high mannose, complex, and hybrid type oligosaccharides. All *N*-linked oligosaccharides contain the pentasaccharide Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAc as a common core (the "trimannosyl core"). High-mannose type sugars contain only α-mannosyl residues in addition to the trimannosyl core, complex type saccharides contain no mannose residues other than those in the trimannosyl core and the chains normally terminate with versions of the Neu5Acα2-3/6Galβ1-4GlcNAcβ1-R (R = trimannosylcore) moiety, and hybrid-type have both complex and high mannose type characteristics.

1.2 Carbohydrates Associated with Disease

1.2.1 Inflammation

The high-fidelity of carbohydrate binding proteins is crucial for the health of an organism. However, this specificity may also be disadvantageous if the carbohydrate recognition is unregulated. The recognition of oligosaccharides produced on the surface of endothelial cells by circulating leucocytes is essential for the correct function of the immune system. Injured or infected cells defensively secret cytokines. Cytokines stimulate endothelial cells to express lectins and display them on their surfaces. These lectins, often called selectins or LEC-CAMs (leucocyte-cell (or lectin) adhesion molecules), have differing affinities for an oligosaccharide present on the surface of passing white blood cells (sialyl Lewis ^{x}, sLe^{x}, see Figure 1.4). Leucocytes are thus snared by endothelial

cells expressing these selectins and once attached, a leucocyte can leave the bloodstream by squeezing through adjacent endothelial cells.

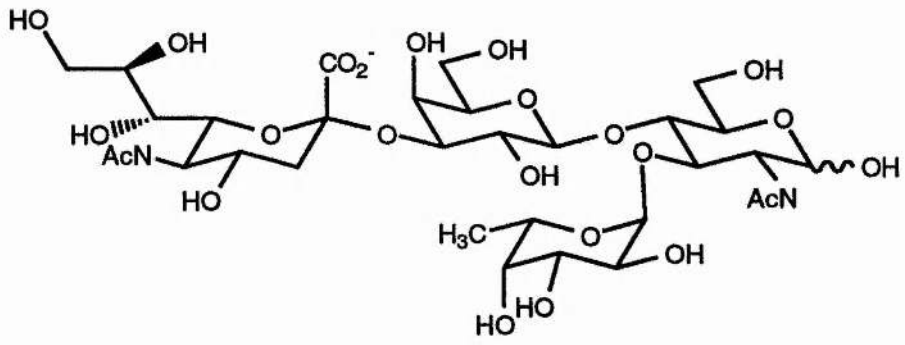


Figure 1.4: Structure of the sialyl Lewis ^{x} antigen (sLe^x), (Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAc.

However, this same vital mechanism also allows leucocytes to accumulate in tissues causing tissue damage, swelling and pain. Certain forms of rheumatoid arthritis (RA) and thrombosis are caused by the accumulation of white blood cells. Reperfusion injury is a disorder that occurs after the flow of blood is temporarily cut off from tissue, such as during heart attack, when the blood flow resumes, leucocytes destroy tissues damaged by lack of oxygen leading to necrosis and dangerous swelling.

1.2.2 Oncology

Changes in the glycosylation profiles of cells is often associated with disease and is a manifestation of defects in sugar metabolic pathways caused by cells undergoing oncogenesis, for example. Certain cancerous cell lines (reviewed in (Oettgen, 1989)) have higher densities of simpler gangliosides such as G_{M3} (Neu5Acα2-3Galβ1-4GlcCer) and G_{D3} (Neu5Acα2-8Neu5Acα2-3Galβ1-4GlcCer) at the expense of more complex gangliosides such as G_{M1} (*See Figure 1.2*). Monoclonal antibodies (mAbs)

raised against theses sugars or plant lectins which recognise the G_{M3}/G_{D3} epitopes are often used in the clinical diagnosis of cancer cell-lines. Sialyl Lewis ^{x} (*See Figure 1.4*) has also been shown to play a role in tumour metastasis. Expression of sLe^{x} on the surface of cancer cells promotes metastasis by providing adhesion and invasion points for circulating metastatic cells.

1.3 Microbial Recognition of Cellular Carbohydrates

With the vast abundance of oligosaccharides expressed on the surface of animal cells, it is not surprising that microbes have 'hijacked' cell-surface oligosaccharides to promote their own adhesion and infection (Reviewed in (Sharon and Halina, 1990; Karlsson *et al.* 1992; Lingwood, 1992; Sharon and Lis, 1993; Karlsson, 1995)). Indeed, the current complexity of biological glycoconjugates may be the result of an evolutionary battle between microorganisms and their hosts. Bacteria (and their toxins), viruses, and protozoa all produce lectins. Adhesion is a prerequisite of infection as the body's normal cleaning mechanisms would clear microbes, for instance, from the respiratory tract or the urinary tract. Microbes must therefore be able to bind selectively to tissues they require for proliferation. The ability of carbohydrates to block hemagglutination by bacterial lectins was first noted by J. P. Duguild in Dundee in the 1950's. *In vivo* it was found that alpha-mannoside could prevent the binding of a mannose-specific strain of *Escherichia coli* to the bladder of mice, thus preventing infection. Similarly the clearance of P-fimbriated *E. coli* from murine kidneys could be enhanced by administering globoside (Gb_4; *Figure 1.2*) to the mice. Trials involving the clinical treatment of *E. coli*

K99 infected calves with free oligosaccharides (glycolipids with the fatty acid tail removed) isolated from cow plasma successfully cured these animals of potentially fatal diarrhoea (de Graaf *et al.* 1980). These observations first highlighted the possibility that carbohydrates or drugs based on the natural carbohydrate receptor may be helpful in the clinical treatment of bacterial and viral disease.

1.3.1 Bacterial Toxins

Certain pathogenic strains of bacteria produce toxins which also recognise cell-surface glycoconjugates. Recognition of specific cellular oligosaccharides is essential for toxin invasion and ultimately cellular death. The 'AB_5' toxins (a subset of the 'A-B' family of toxins) are a group of structurally related proteins produced by a variety of bacteria. The *Vibrio cholerae* cholera toxin (CT), *E. coli* heat labile enterotoxin (LT) and verotoxins (VTs) (also known as shiga-like toxins), *Shigella dysenteriae* shiga toxin (ST), and *Bordetella pertussis* pertussis toxin (PT) all share the AB_5 topology but their modes of cytotoxicity are very different. AB_5 toxins consist of a single catalytic A ('active') protomer and a pentameric B ('binding') oligomer (*See Figure 1.5A*) which binds specific oligosaccharide motifs; the B subunit may either be homopentameric (CT, LT, ST, and VTs) or heteropentameric (PT). The A subunit is required for cytotoxicity, however, it remains inactive without its B subunit partner to initiate cellular invasion. A remarkable feature of the B subunits in the AB_5 family is the high structural homology these protein posses despite the distinct lack of sequence homology between CT/LT and VT/ST, and PT families.

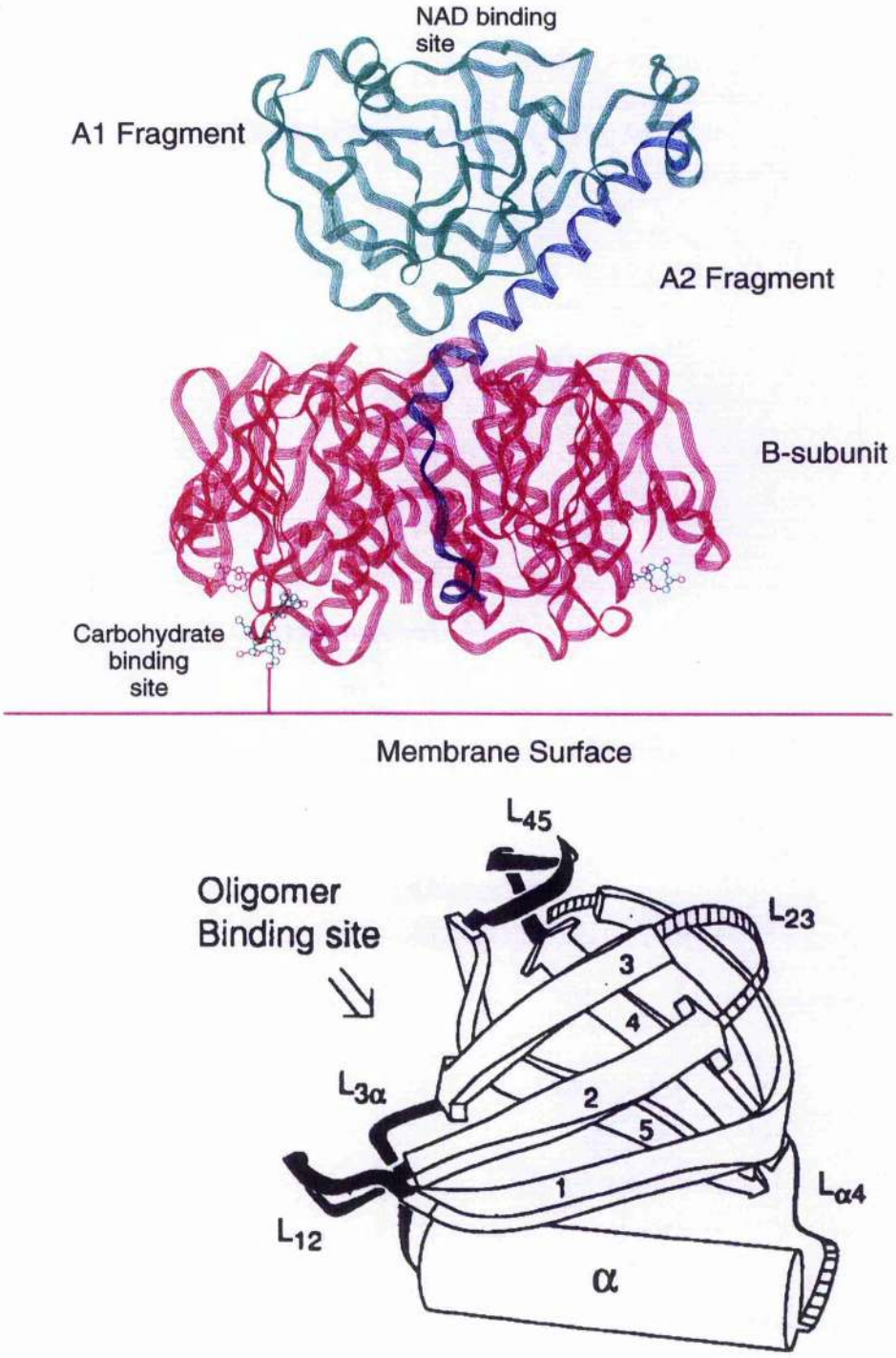


Figure 1.5: (Above) A schematic representation of the cholera toxin highlighting the key structural points. (Below) A schematic representation of the OB fold (Murzin, A. G. 1993) of the B-subunit shared by all AB_5 toxins.

The B-subunits possess the same fold: six anti-parallel β-strands forming a closed β-barrel, capped by an α-helix between the fourth and fifth strands (known as the OB (oligomer binding) fold)(Murzin, A. G. 1993) (*See Figure 1.5B*). These monomers arrange themselves into a 'torus' configuration (*See Figure 1.6*). The A2 fragment of the A subunit is threaded through this pore and this interaction is believed to hold the toxin together in the extra-cellular environment. The pores themselves posses strikingly different chemical properties. The central pore of LT and CT are primarily lined with charged amino acids, whilst the central pore of VT and ST is hydrophobic in nature (Sixma, T. K. *et al.* 1993). What effects the differences in chemical environment has on the possible translocation of the A subunit across the cellular membrane remains unclear.

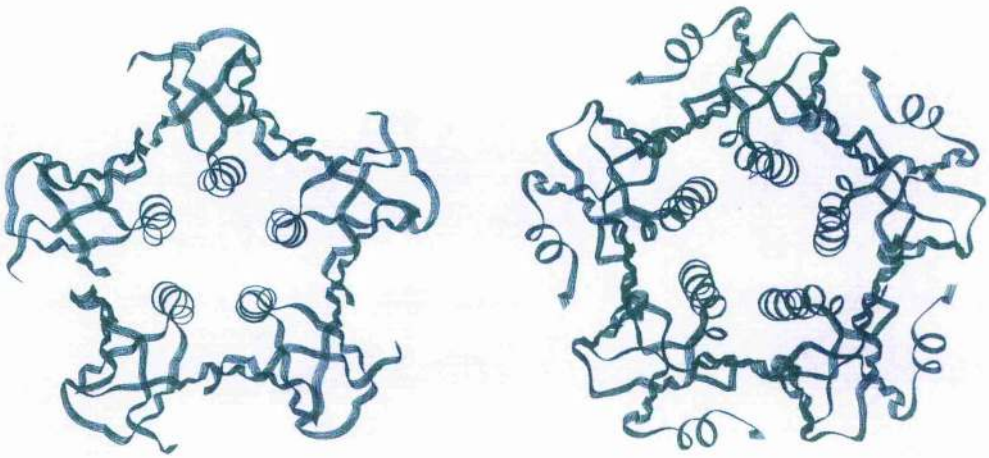


Figure 1.6: A ribbon representation of the B-subunits of the (left) vero toxin and (right) cholera toxin.

1.3.2 Structure and Function of Cholera Toxin and the related E. coli Heat-Labile Toxin.

Cholera remains an important problem in the Third World, killing millions of people annually in epidemics covering much of South America and Southern Asia (For an extensive review the reader is directed to (Spangler, 1992; Burnette, 1994; and Hol *et al.* 1995)). The disease is characterised by severe dehydration and cramps symptomatic of the rapid, extreme loss of fluid and electrolytes during the course of infection. The diarrhoea is caused by the action of the cholera toxin (CT) which recognises a specific glycosphingolipid (G_{M1}; *Figure 1.2*) present on the surface of intestinal epithelial cells (Fishman 1978; Fishman 1982a). The structurally related *E. coli* heat labile enterotoxin causes a much milder disease in humans, known as traveller's diarrhoea also recognises G_{M1} as well as a neutral neo-lacto sugar (lacto-*N*-neo tetraose; *Figure 1.2*) (Griffiths, 1986; Teneberg *et al.* 1994).

The A subunit of these toxins destroys cells in an identical manner: once the B subunit has bound to the cell-surface, the whole toxin or just the A subunit is translocated across the membrane (the exact process has not been elucidated); the A subunit then enters the cytosol where it is cleaved into two fragments A1 and A2. The activated A1 fragment is capable of binding NAD and catalysing the ADP-ribosylation of $G_{s\alpha}$, a GTP-binding regulatory protein associated with adenylate cyclase. Activation of this protein by binding GTP stimulates the cyclase system, hydrolysis of GTP reverts the cyclase system back to its inactive state. ADP-ribosylation of $G_{s\alpha}$ inhibits the hydrolysis of GTP and the adenylate cyclase system remains activated. This and other factors such as an increase prostaglandin synthesis, causes the excessive accumulation of salt and

water in the intestinal lumen and ultimately cell death (Fishman 1980b; Janico, 1991; Burnette, 1994).

The B subunits of CT and LT share high sequence and structural homologies (> 80%) and high resolution X-ray crystal structures of CT in complex with oligo-G_{M1} (G_{M1} lacking the ceramide portion) (Merritt 1994) and LT in complex with lactose (Sixma *et al.* 1992) has been solved. These provide a molecular basis for the specificity of these two toxins to G_{M1}. Analysis of the binding sites of CT and LT highlight some of the common features of carbohydrate binding sites (Reviewed in (Quiocho *et al.* 1989a,b) (*Shown in Figure 1.7*): the specificity for the terminal sugar (galactose) stems from the fact that almost all hydrogen bond donors and acceptors of galactose are engaged in hydrogen bonding. There are also two well defined water molecules which mediate H-bonds between galactose and the protein. A prominent feature of carbohydrate sites is 'stacking' of sugars against aromatic rings. In the CT/LT galactose binding site there are extensive van der Waals contacts between the relatively hydrophobic B face of galactose and the aromatic ring of Trp88 and the diastereotopic hydrogens of the galactose hydroxymethyl group stack against histidine 57. The orientation of the aromatic ring of Trp88 also offers ensures specificity for galactose as it prevents the binding of the C4 epimer of galactose - glucose: this would result in a steric clash between the glucose 4-OH group and the plane of the ring. These features are common to almost all carbohydrate recognition domains studied so-far. Unlike other Neu5Ac binding proteins, the sialic acid moiety of G_{M1} makes little intimate contact with the protein and the interaction between the carboxylate group and the protein is mediated largely through a bound water. This was a surprising observation as any chemical

modifications of G_{M1} at the carboxylate group resulted in a decrease in binding affinity (Schengrund and Ringler, 1989). Both LT and CT toxins bind to G_{M1} very strongly, the equilibrium dissociation constant (K_D) for G_{M1} is in the nanomolar range. The binding affinity for oligo-G_{M1} is still in the micromolar range, in contrast, the typical K_D values for most lectins interacting with isolated oligosaccharides is typically in the millimolar range (see for example, the verotoxins and plant lectins). Thus, there may be some scope in developing G_{M1} based drugs to inhibit the cellular attachment of CT or LT.

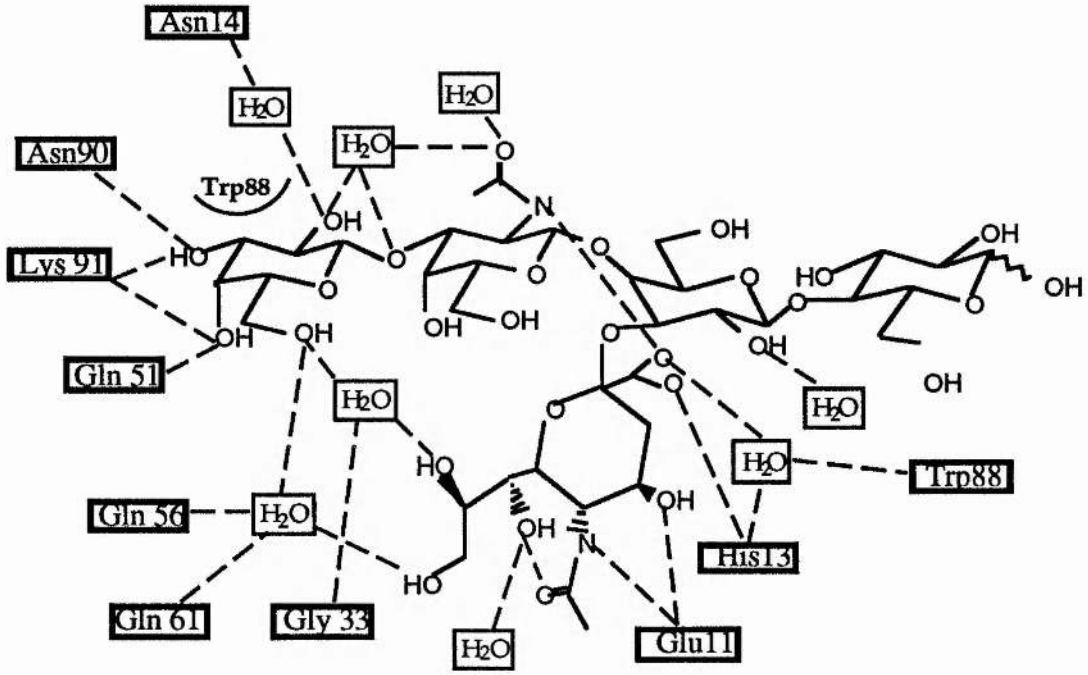


Figure 1.7: Schematic representation of the cholera toxin binding site. Dashed lines represent H-bonds. The stacking interaction between the B-face of the terminal galactose moiety and tryptophan 88 is also highlighted.

1.3.3 Structure and Function of Enteropathogenic E.coli Verotoxins

Verotoxins (VTs) are elaborated by certain strains of *Escherichia coli*. (O'Brien and Holmes, 1987; Jackson, 1990). Infection with VT producing *E. coli* is associated with hemolytic uremic syndrome (HUS), which is the major cause of paediatric acute renal failure. The toxins also cause diarrhoea or bloody diarrhoea (haemorrhagic colitis). Whereas, CT and LT related diseases are common in the Third World, infection with VT-producing *E. coli*, most infamously serotype O157:H7 has caused fatal outbreaks of food poisoning in developed countries such as the United States of America and the United Kingdom. Cytotoxicity arises from the ability of the A subunit to inhibit protein synthesis by cleaving A4324 from 28S rRNA, thus inactivating mammalian ribosomes by preventing the EF-1-dependent binding of aminoacyl tRNA to the 60S ribosomal subunit. Similar N-glycosidase activity has been reported for the A subunit of the highly toxic plant lectin from *Ricinus communis* (ricin toxin), its self a member of the A-B toxin family (Olsen and Phil, 1982).

Cells expressing members of the globo-series (*Figure 1.2*) glycolipids have been shown to be susceptible to verotoxin invasion. Verotoxins 1 and 2 (VT1 and VT2) and the shiga toxin are associated with human diseases specifically interact with globotriaosylceramide (Gb_3: Galα1-4Galβ1-4GlcCer) which is expressed on the surface of the paediatric renal glomerulus and gastrointestinal tract (Boyd *et al.* 1993). VT2e, which is associated with edema disease of swine, binds preferentially to globotetraosylceramide (globoside, Gb_4, shown in *Figure 1.2*) (DeGrandis *et al.* 1989). The B-subunit of VT1 has been solved by X-ray crystallography (Stein, P. E. *et al.* 1992), however, the oligosaccharide

binding sites were not readily identified. Solubility problems prevented the co-crystallisation of VT B-subunit with Gb_3, and the toxin binds oligo-Gb_3 very weakly (K_D ~10^{-3} M for oligo-Gb_3, as compared to ~10^{-9} M for intact Gb_3). Site directed mutagenesis studies identified several amino acids which were essential for glycolipid attachment (Boyd *et al.* 1993), for example, the binding phenotype of the toxins could be shifted from Gb_3 to Gb_4 by switching Asp18 to Asn (Tyrrel, 1992). A putative binding site situated in a cleft between the B subunit monomers was proposed on the basis of these studies. Molecular modelling investigations based on molecular affinity potentials (Goodford, 1985) performed on the X-ray crystal structure identified two possible binding sites per monomer. In common with the G_{M1}-CT complex, Gb_3 is believed to form a hydrophobic interaction with an aromatic ring, in this case phenylalanine 30 (Nyholm *et al.* 1995). A possible inhibitor of all the verotoxins was proposed on the basis of the modelling studies. AminoGb_4, a de-*N*-acetylated derivative of Gb_4, was predicted to form a salt-bridge to a group of conserved aspartates (Nyholm*et al.* 1996a,b). Inhibition studies confirmed that aminoGb_4 can bind to all the verotoxins, thus, it may be possible to design potent inhibitors of VT attachment based on the aminoGb_4 structure.

The structure of the verotoxin in free solution has been solved by high resolution multidimensional heteronuclear NMR spectroscopy (Richardson*et al.* 1997). The observed asymmetry of the B-subunit by X-ray crystallography was not reported in free solution. The weak interaction between oligo-Gb_3 and VT, which prevented the identification of the oligosaccharide binding site by X-ray crystallography, was an

advantage in NMR studies as it allowed the measurement of transferred nuclear Overhauser enhancements (TRNOE) (Clore and Gronenborn, 1982). These studies confirmed that the theoretical bound conformation of Gb_3 was correct. The measurement of the effects aromatic rings have on chemical shifts (ring current shifts, (Perkins, 1982)) also confirmed the stacking of the sugar against Phe30.

1.4 Plant Lectins as Models for Oligosaccharide-Protein Interactions

Plants express large quantities of lectins (often up to 40% of the total protein in some plant tissues), however, their biological function remains obscure. Several hypothesis have been put forward for the role of these proteins including a role in plant self-defence: lectins often have some antifungal properties and the ability of some of plant lectins to bind sialic acids (sugars only produced in large quantities by animals and bacteria) also adds weight to this argument. Lectins may also promote the symbiotic association of nitrogen fixing microorganisms such as *rhizobium* to leguminous plants.

The ability of plant lectins to bind reversibly particular carbohydrates has made the purification of lectins, by affinity chromatography on immobilised carbohydrates, trivial and several hundred lectins have been isolated to date. Their specificity, ease of preparation coupled with the large quantities of lectin expressed per gram of plant biomass, make lectins excellent tools in the fields of glycobiology, histology, hematology, rheumatology, and oncology. The ability of plant lectins to agglutinate blood cells was first noted in the 1900's and since then lectins have been used extensively in the typing of blood products. For example, lectins

from *Lotous tertragonolobus* and *Ulex europaeus* serve to identify blood type O cells. In surgery, soybean agglutinin is routinely used in the removal of immunocompetent lymphocytes from human bone marrow prior to transplantation. Plant lectins are also used to study changes in the glycosylation patterns of intact cells as well as individual proteins and lipids. Malignant cells, for example, will bind to a different range of lectins compared to their healthy counterparts. Lectins may also be used to determine the invasive potential of tumours.

Plant lectins have proven particularly attractive model compounds for the study of oligosaccharide - protein interactions by X-ray crystallography, microcalorimetry, and NMR spectroscopy. Typical examples of the binding site architecture highlighted in the CT discussion where first observed in the X-ray crystal structures plant lectins in complex with mono- or small oligosaccharides (Reviewed in (Einsphar, 1989; Quiocho *et al.* 1989a; Rutenber and Robertus 1989; Lis and Sharon, 1991; Vyas 1991; Bundle, 1992; Spurlino *et al.* 1992; Naismith & Field, 1996).

1.5 Carbohydrate Therapeutics

It is clear from the few examples discussed above that the prevention of carbohydrate recognition may have potentially important therapeutic benefits. Carbohydrate based drugs may also have other advantages such as low toxicity and immunogenicity. In addition, in the treatment of microbial infections, these drugs merely prevent the adhesion of microbes and not their death, there is therefore less evolutionary pressure for mutation, so resistance is less likely. There are currently some fifteen carbohydrate drugs undergoing clinical trials for ailments including

reperfusion injury, cancer, diarrhoea, and also thrombosis, diabetes, ulcers, epilepsy, and Parkinson's disease ((McAuliffe and Hindsgaul, 1997) and references therein).

Carbohydrate drugs may be designed to work in several ways:

1. *Inhibition of protein attachment*

The most common method so far is to design a drug, based on the natural carbohydrate ligand, which would bind irreversibly to the pathogenic protein, thus preventing cellular adhesion. This method has been used in the treatment of reperfusion injury with both Cytel and Glaxo Wellcome currently designing small sLe^x mimics to prevent the accumulation of leucocytes in tissues damaged by heart attacks or surgery. The clearance of bacteria and their toxins from the gastrointestinal tract with oligosaccharides has long been used in nature: mammalian milk, which contains high concentrations of oligosaccharides often found on the surface of the gut, is believed to protect infants from infections. Man-made attempts to mimic this process include the development of a Gb_3 conjugate linked to an insoluble powder to prevent verotoxin adhesion.

2. *Alteration of carbohydrate biosynthesis*

Many diseases are associated with defects in oligosaccharide metabolism. Several azasugars (nitrogen analogues of carbohydrates) may help in the treatment of certain forms of diabetes by inhibiting sucrase and amylase (responsible for the breakdown of polysaccharides in the gut) and therefore prevent the adsorption of glucose into the bloodstream. The spread of metastasic cells may be prevented by inhibiting the synthesis of

glycoconjugates such as sLe^x known to aid the colonisation of healthy tissues by cancerous cells. Inhibiting the catabolism of Neu5Ac terminating oligosaccharides has been used successfully in the treatment of influenza (von Itzstein *et al.* 1993).

3. *Eliciting an immune response*

As discussed previously, cells undergoing oncogenesis often express a different range of glycoconjugates compared to their healthy counterparts. The treatment of cancer patients by stimulating anti-ganglioside antibodies has had limited success (Livingston *et al.* 1987). Recently, a similar methodology has been applied successfully to the treatment of lung, breast, colon, and ovarian cancers: this treatment consists of injecting the patient with a glycopeptide (found on the surface of these cancer cell lines) hapten which elicits an immune response against these cells.

An interesting development in the design of therapeutics is the application of the carbohydrate binding toxins such as verotoxin and cholera toxin as selective cytotoxins and drug delivery systems. Killing tumours which express high levels of Gb_3 with verotoxins has been proposed (Arab *et al.* 1995). The cholera toxin and cholera toxin heat stabile toxin adducts have been used as an antigen delivery system in order to enhance the immunogenicity of haptens and may also perform a role in vaccine design (Aitken and Hirst, 1995; Sanchez *et al.* 1997).

1.5.1 Rational Drug Design

A prerequisite for a detailed understanding of the molecular basis of the interaction of a carbohydrate with its protein receptor is a high resolution three dimensional structure of the complex. Several important crystallographic studies have been described recently which illustrate in detail the precise nature of certain carbohydrate-protein interactions in the solid state. Information obtained from these studies may be used in the directed synthesis of novel drugs using the natural carbohydrate moiety as a lead compound. Perhaps the most celebrated example of model based oligosaccharide drug design was the seminal work performed by von Izstein's laboratory on the development of an influenza sialidase inhibitor (von Itzstein *et al.* 1993). Sialidase helps the elution of newly synthesised influenza virons from infected cells by cleaving terminal Neu5Ac residues from host cell-surface glycans, and it also assists the movement of the virus through the mucus within the respiratory tract. Inhibition of this enzyme results in limiting the establishment and progression of infection. Computer-aided molecular modelling was used to analyse the active site and predict functional group changes which could be made to sialic acid in order to increase its binding affinity. Modelling studies predicted that substitution of the 4 hydroxyl group with an amino or guanidinyl group should produce a significant increase in the overall binding. Results show that the 4-guanidino derivative has an inhibition constant (K_i) for various strains of influenza in the ~10^{-10} range. In addition the affinity of this compound against human sialidases is a million fold lower, therefore offering excellent selectivity for virus sialidases.

Detailed X-ray crystal structures of biologically important oligosaccharide complexes are, however, rare as the relatively flexible oligosaccharide often results in poor electron densities and carbohydrates are often only observed where they are stabilised by crystal packing forces which may orientate the sugar in an unnatural conformation. Many lectins bind isolated oligosaccharides analogues of their receptor glycans very weakly and thus, may not be co-crystallised (Stein *et al.* 1992). An additional problem in the solid state is that carbohydrate binding sites identified by X-ray crystallography may not be "active" in the less ordered solution state (compare (Wright, 1992) with (Wright and Kellog, 1996)).

In contrast, high resolution structural studies of glycan protein interactions by nuclear magnetic resonance (NMR) spectroscopy can offer complimentary information on the structure of the carbohydrate part of a sugar-protein complex. Solution studies also have important potential advantages since a comparison of the solution structure of the free ligand with that of its bound state is more meaningful, and moreover the dynamics of the system are accessible.

1.5.2 Nuclear Magnetic Resonance Spectroscopy (NMR) of Carbohydrates and Glycoconjugates

NMR has been used to study carbohydrates, since the 1950's, and is perhaps uniquely suited to the investigation of carbohydrates and their complexes. NMR has been used to determine the type, number, and primary sequence of oligosaccharides, including the occurrence of branch points and the location of appended non-carbohydrate groups; and their three dimensional (3D) conformations(s) and dynamics in free solution and in complex with proteins, and orientated in a membrane

environment ((van Halbeek 1996) and references therein).

1.6 NMR Spectroscopy

The nuclear magnetic resonance phenomenon, first observed in 1946, has become a widespread tool for the non-destructive analysis of both organic and inorganic compounds. NMR techniques can provide information on structure, conformation, and internal mobility. Sanders and Hunter (1987) provides an excellent introduction to the subject.

The fundamental basis of the NMR experiment is perturbation by a radiofrequency pulse, of equilibrium magnetisation from an axis parallel with the static magnetic field, B_0, into a vector perpendicular to this axis, precessing with a characteristic (Larmor) frequency (ω). The value of the Larmor frequency is due to the extent to which the nucleus is shielded from the external field by nearby electron-inductive groups.

$$\omega = \gamma B_0 \qquad \textit{Eqn 1.1}$$

1.6.1. Conformational Analysis

The following three parameters are particularly important in the conformational analysis of oligosaccharides in free solution.

(a) Relaxation properties (NOE, T_1, T_2)

(b) Spin-spin coupling constants (J)

The conformation of oligosaccharides in complex with proteins may be determined from the following experiments.

(c) Line-shape analysis

(d) Transferred NOE

Detailed explanations of NMR relaxation phenomena are given by Noggle and Schirmer (1971) Neuhaus and Williamson (1989), and analysis of the transferred NOE are given by (Ni 1994; Ni and Scheraga A. 1994), and (Clore and Gronenborn 1982, 1983a,b).

After application of a pulse, or other perturbation, nuclei relax to their equilibrium states by one of two mechanisms. Longitudinal (T_1) relaxation causes the population difference between two spin states of a given nucleus to return exponentially to equilibrium, due to transfer of energy to the surroundings or 'lattice'. In terms of classical formalism, bulk magnetisation in the X-Y plane returns to the Z axis. The second mechanism, spin-spin (T_2) relaxation causes a loss in phase coherence due to mutual exchange of spin energy, and resultant decay in the bulk intensity of transverse vectors. (Bloch model).

Both mechanisms of relaxation are due to time-dependent magnetic or electric fields, derived from random thermal motions within the sample. For both ^1H and ^{13}C nuclei, the major sources of such fields, the intensity of which is termed the spectral density $J(\omega)$, are the magnetic moments of neighbouring protons. Relaxation is most efficient for both T_1 and T_2 when the timescale of such interactions is at or near the Larmor

frequency. T_2 is additionally sensitive to low-frequency modulations of spin energy levels. Thus the relaxation mechanisms are dominated by the effects of dipole-dipole interactions with adjacent ^1H nuclei. The strength of these dipolar interactions are primarily dependent, among other factors, upon the internuclear distance, and the mobility of the vector within the static B_0 field. Thus these relaxation phenomena can provide information about both geometric and dynamic behaviour.

1.6.2 Nuclear Overhauser Effect

Consider a system of two ^1H spins, I and S, that are not spin coupled, but are close in space.

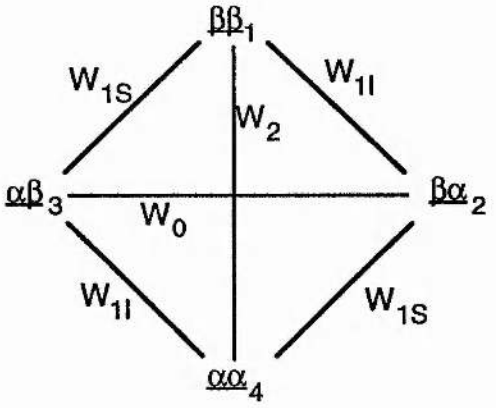


Figure 1.8 : Energy level diagram for two non-coupled spins, I and S, in close spatial proximity. W_1, W_2 and W_0 represent the transition probabilities for single double and zero quantum processes. Spin states are labelled α or β.

We can define transition probabilities between states as shown in figure 1.8. The populations of each energy level are given by the Boltzmann distribution. Application of a radiofrequency pulse at the frequency of S saturates this spin causing the populations across the S spin transitions to be equalised. Relaxation will proceed by the various pathways, including

W_2 and W_0 double- and zero-quantum spin transitions. W_2 transitions are promoted by magnetic fields fluctuating at ~2ω, and W_0 by low frequency fluctuations. Hence in slowly tumbling molecules, W_0 processes dominate, causing a reduction in the intensity of the I spin, resulting in a -100% NOE in the limit. In a rapidly tumbling molecule, W_2 predominates, causing a positive enhancement. When the energy level populations are disturbed from equilibrium, the I spin also relaxes by transferring energy to its neighbouring protons. These indirect enhancements are also negative, and after prolonged irradiation of S the transfer of indirect NOE enhancements ('spin-diffusion') causes saturation of all spins in the molecule, and hence the distance proportionality of the NOE magnitude is lost. The rate of cross-relaxation, σ, proportional to r^{-6}, is a more convenient parameter to use than the enhancement itself.

In a multispin system, the rate of intensity change for spin I is:

$$\frac{dI_Z}{dt} = -R_I(I_Z-I_Z^0)-\sigma_{IS}(S_Z-S_Z^0)-\Sigma\sigma_{IM}(M_Z-M_Z^0) \qquad \textit{Eqn 1.2}$$

where I_z and I_z^0 are I spin intensities along the Z-axis at time zero and time t respectively; R_I is the relaxation rate of I; and M represents all other spins in the system. The initial rate, when $I_z=I_z^0$, $M_z=M_z^0$, and $S_z=0$, is given by

$$\frac{dI_z}{dt} = \sigma_{IS}S_z^0 \qquad \textit{Eqn 1.3}$$

For homonuclear ^1H-^1H interactions, σ is given by:

$$\sigma_{is} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_i^2 \gamma_s^2 \hbar^2}{10 r_{is}^6} \left[6J(\omega_i + \omega_s) - J(\omega_i - \omega_s) \right] \qquad \textit{Eqn 1.4}$$

Where $\hbar$ is Planck's constant divided by 2π, μ_0 is the permeability of free space, r_{IS} is the internuclear distance, and γ is the magnetogyric ratio of a proton. $J(\omega)$, the spectral density function, is defined as:

$$J(\omega) = \frac{\tau_c}{1+\omega^2 \tau_c^2} \qquad \textit{Eqn 1.5}$$

τ_c (the correlation time for molecular reorientation) is inversely related to rates of molecular motions, and is identical for each ^1H-^1H vector in a rigid isotropically tumbling molecule.

Using the initial rate approximation, immediately after saturation of S, the NOE at I is affected by only one distance-dependent term, σ_{IS}. Therefore, in such situations, all enhancements behave as though they were in a two-spin system, and those measured during truncated driven experiments, when the saturation period is within the linear build-up region, are in proportion to the internuclear separation from S. With longer presaturation times, indirect enhancements develop, and distance specificity is lost.

When all constant terms in equation are gathered together as a single constant k, the equation simplifies to:

$$\sigma = kr^{-6} \qquad \textit{Eqn 1.6}$$

Then, when an NOE enhancement is observed between two protons, (I and M), and the distance between I and S is known, the S-M distance can be estimated from a simple ratio calculation:

$$r_{sm}=r_{is}(\sigma_{is}/\sigma_{sm})^{1/6} \qquad \textit{Eqn 1.7}$$

Due to the r^{-6} dependence of small inaccuracies in measured NOE enhancements have negligible effects in calculated internuclear distances. The reference distance method also referred to as the isolated spin-pair approximation (ISPA) method may be used to model theoretical NOE intensities from static or dynamic structures providing the following caveats are appreciated:

1. The accuracy of the calculated S-M distance relies upon that of the reference distance.
2. Inaccuracies are introduced by non-instantaneous saturation.
3. The method is only valid for the initial rate approximation.
4. Integration of spectra, particularly with overlapping or low intensity signals can be inaccurate.
5. Internuclear vectors connecting IS and SM must have the same effective correlation time (τ_e). This will not be the case if the molecule exhibits anisotropic tumbling or has flexibility in the S-M distance.
6. In the event of internal motions, enhancements are heavily weighted by the conformations with closest contact, due to the r^{-6} dependence. It is the internuclear distance, rather than the

enhancement itself which is averaged.

7. Nuclei I, S, and M must have the same γ value and must not be relaxed by a directly attached NMR 'active' nucleus.
8. Nuclei I, S, and M must not be undergoing chemical exchange.

Providing these conditions can be satisfied, the S-M distance can be calculated to within ~10%. However, in flexible polysaccharides (5 and 6 above), inequalities of τ_c become relevant. Internal mobility adds uncertainties to the proportionality of σ and r^{-6} (Genest, 1989). For a flexible molecule in multisite conformational exchange that is slow on the τ_c timescale, the effective internuclear distance is simply a time-average of the separation at each individual conformation $\langle r^{-6} \rangle^{-1/6}$.

1.6.3 Full Relaxation Matrix Calculations

In contrast to the ISPA method, full relaxation matrix (FRM) calculations offer a more theoretically rigorous method for the interpretation of (R)NOESY data. The aim of FRM calculations is to determine the longitudinal relaxation of a complete spin-system, thus, the effects of spin-diffusion, the presence of heteronuclei, and chemical exchange may be calculated from a fixed or dynamic model of the system if a suitably parametrised algorithm is available. The full relaxation rate matrix **R** is calculated (assuming cross-correlation can be ignored) by solving a set of generalised Bloch equations, which govern the time dependence of the magnetisation M_{zI} of each spin I within a molecule. These coupled differential equations are best calculated in the form of a matrix:

$$\frac{d}{dt}\begin{bmatrix} M_I \\ M_S \\ \vdots \\ M_n \end{bmatrix} = -\begin{bmatrix} R_{II} & R_{IS} & \cdots & R_{In} \\ R_{SI} & R_{SS} & \cdots & R_{Sn} \\ \vdots & \vdots & & \vdots \\ R_{n1} & R_{n2} & \cdots & R_{nn} \end{bmatrix}\begin{bmatrix} M_I \\ M_S \\ \vdots \\ M_n \end{bmatrix} \qquad \textit{Eqn 1.8}$$

with $M_I = M_{zI} - M_{zI0}$. Here M_{zI0} is the z magnetisation of spin I at equilibrium. R_{II} is given by

$$R_{II} = \sum_S \rho_{IS} + \rho_I^* \qquad \textit{Eqn 1.9}$$

the dipolar contributions to the spin-lattice relaxation rate ρ_{IS} are given by

$$\rho_{IS} = \frac{\gamma_I^2 \gamma_S^2 \hbar^2}{10 r_{IS}^6} \sum_{S \neq I} [J_0(\omega) + 3J_1(\omega) + 6J_2(2\omega)] \qquad \textit{Eqn 1.10}$$

and ρ_i^* represents other (nondipolar) relaxation sinks ("leakage") in macromolecule systems, and the cross-relaxation rate R_{IS} is equal to σ_{IS} (*See equation 1.4*).

If all the interproton vectors are assumed to be rigid and to move isotropically, the spectral density function takes the form given in equation 1.5.

Equation 1.8 is rewritten as:

$$\frac{d}{dt}\mathbf{M} = -\mathbf{R}\mathbf{M} \qquad \textit{Eqn 1.11}$$

where the **R** is the relaxation rate matrix. The solution of the set of generalised Bloch equations is then:

$$\mathbf{M}(\tau_M) = e^{(-\mathbf{R}\tau_m)}\mathbf{M}(0) \qquad \textit{Eqn 1.12}$$

These equations can be generalised for the analysis of 2D NOESY spectra in which case $\mathbf{M}(\tau_M)$ is a matrix proportional to the normalised NOE intensities recorded with a mixing time τ_M and $\mathbf{M}(0)$ is the diagonal matrix containing the z magnetisation.

1.6.4 Spin-coupling Constants

The degree of atomic orbital overlap affects the magnitude of spin-coupling constants, which are therefore related to the dihedral angle (θ) between vicinal coupled spins (Karplus, 1959; Karplus, 1963). The generalised Karplus relationship is applicable to both homonuclear and heteronuclear spin-coupling constants, and is of the form:

$$J = A\cos^2\theta + B\cos\theta + C \qquad \textit{Eqn 1.13}$$

Where A, B and C are constants for which different values have been proposed depending on the torsional angles under investigation. Haasnoot has included additional terms for substituent electronegativities (Haasnoot *et. al.*, 1980). The ^{1}H-^{1}H spin-coupling constants and chemical shifts of pro-R and pro-S hydroxymethyl resonances have been assigned unequivocally for several hexapyranosides (Nishida *et al.*, 1987). H-6_{proR} invariably has a greater chemical shift and larger H5-H6 coupling constant than H-6_{proS}. The rotomer distribution around the C5-C6 bond can be found from analysis of the ^{1}H-^{1}H spin-coupling constants.

1.6.5 Two and Three Dimensional NMR Spectroscopy

Resonance assignment, even of simple oligosaccharides, is complicated by the tendency for the majority of resonances in both ^{1}H and ^{13}C spectra to lie within an unresolved envelope, spanning only a few hundred Hertz. Anomeric proton or carbon resonances are distinct from this envelope, being well resolved at low field, due to the electron-withdrawing ring oxygen. Problems of coincident resonances are partially resolved by separation into a second or third dimension, increasing spectral dispersion. The precise nature of the information contained in the spectrum depends upon the details of the pulse sequence and phase cycling used, however a number of key points are shared by all multidimensional experiments. All nD spectra consist of 'building blocks' (Edison *et al.* 1994), consisting of a sequence of pulses to generate the desired coherences, which then evolve under a free precession Hamiltonian during an incremented delay (t_1). In 3D experiments further building blocks transfer magnetisation to other spins which may be recorded during an additional incremented delay. The signal is then detected directly during acquisition. Two dimensional spectra can then be processed by application of two orthogonal Fourier transformations. Three dimensional spectra require a third orthogonal transformation.

1.6.6 Heteronuclei Spectral Editing and Filtration

Despite the poor spectral dispersion exhibited in the ^{1}H and ^{13}C NMR spectra of oligosaccharides, it is unlikely that any two carbohydrate ring protons with the same chemical shift would have an attached carbon nuclei with identical carbon chemical shifts. Therefore correlating proton resonances into a third dimension by their directly attached carbon nuclei

greatly simplifies the analysis of traditional 2D experiments such as ^{1}H-^{1}H ROESY or NOESY (*Figure 1.9*) The assignment of oligosaccharides may also be aided through the acquisition of heteronuclear analogues of COSY and TOCSY experiments known as HCCH-COSY and HCCH-TOCSY, respectively (Bax *et al.* 1990). The advantages of acquiring HCCH-COSY and HCCH-TOCSY are two-fold: it allows the simultaneous assignment of both the proton and carbon resonances; and one bond carbon-carbon coupling constants are much larger and more uniform than three-bond proton-proton coupling constants in sugar residues. Therefore, small proton couplings (i.e. between Gal H4 and H5 $^{3}J_{HH}$ ~1 Hz) which do not give rise to an observable cross-peak in 2D COSY/TOCSY spectra, may be mapped in HCCH experiments.

The analysis of protein-ligand complexes by NMR is complicated by the fact that ligand resonances often have coincident shifts with the protein spectral envelope and are swamped by the more numerous protein signals. However, stable isotope enrichment of the ligand or the protein with one or more heteronuclei (typically ^{13}C and ^{15}N) allows the separation of proton NMR resonances into various subspectra by isotope editing (mapping protons attached to heteronuclei) or filtration (eliminating signals from heteronuclei coupled protons leaving only protons attached to carbon-12 in the final spectrum) (*See Figure 1.9*).

The cornerstone of heteronuclear NMR experiments is the INEPT (Insensitive Nuclei Enhanced by Polarisation Transfer) and reverse INEPT building blocks (Morris and Freeman 1979). The purpose of the INEPT is to transfer in-phase I magnetisation (created after the first ^{1}H 90°

pulse) to anti-phase S nuclei magnetisation on the directly coupled heteronuclei.

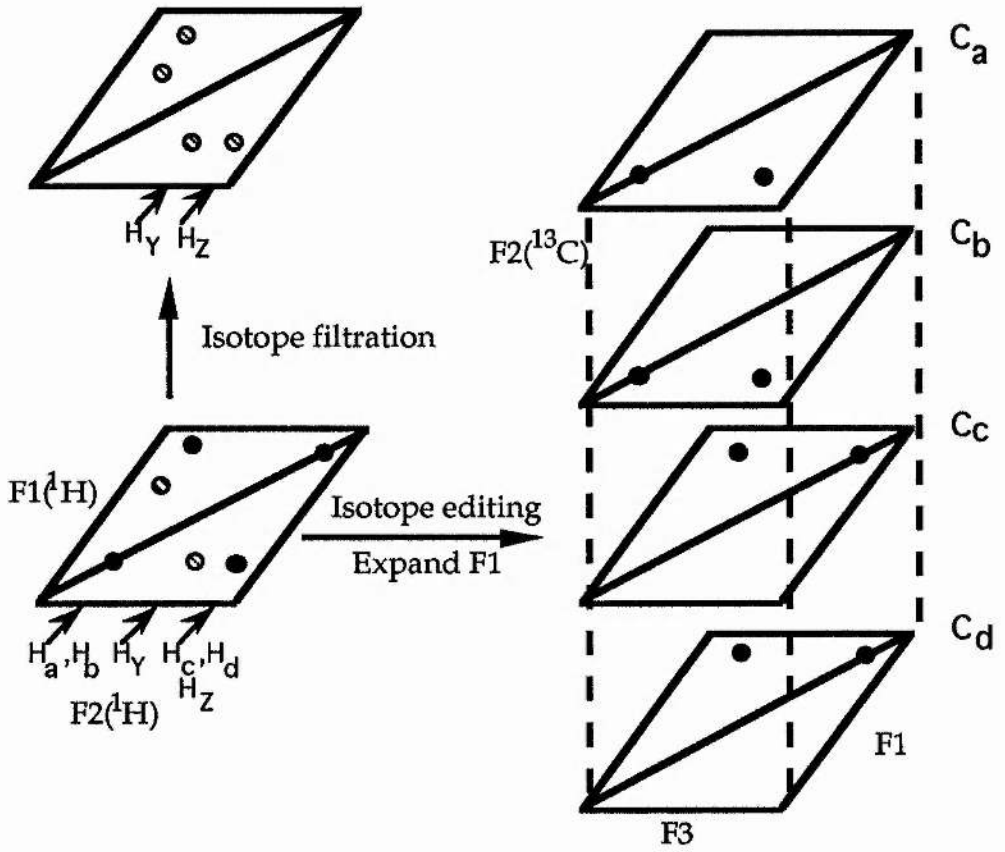


Figure 1.9: Schematic illustration of 2D and 3D NOE spectra of ^{13}C labelled ligand protons with NOEs between H_a-H_c, H_b-H_d, and H_c-H_d and a 'contaminating' unlabelled protein NOE between H_y and H_z. The ^{1}H-^{1}H NOE data of the 2D spectrum are expanded into a third NOE spectrum by frequency labelling the CH groups detected in the F1 (^{1}H) domain by their ^{13}C-resonance frequencies (F2). The 3D experiment is ^{13}C-edited and protein NOEs are not detected, however, cross-peaks due to ^{13}C-^{1}H correlations may be filtered out leaving only protein (^{12}C) NOEs in the top spectrum.

In a two-spin AX system, such as ^{13}C-^{1}H, there are four times the population of the ground state for protons as there for ^{13}C, because the protons have four times the Larmor frequency (because of the ration of the gyromagnetic ratios, γ) and therefore the energy of the transitions is four times greater. By transferring this magnetisation to the carbon-13 nuclei, the detected signal from the carbon will be stronger. The intensity ratio between an INEPT experiment and a simple 90° pulse on carbon is given by:

$$\frac{\text{INEPT observation}}{\text{90° carbon}} = \frac{\gamma I}{\gamma S} \qquad \textit{Eqn 1.14}$$

The intensity of the signal is therefore enhanced by γ_I/γ_S. Additional advantages in using polarisation transfer are: the receiver coil is more sensitive to signals of higher frequency (Cavanagh *et. al.*, 1988), thus ^{1}H signals which resonate at a higher frequency than ^{13}C nuclei in the same magnetic field, will be detected with greater sensitivity. Secondly, signal repetition can be on the order of T_1 for the sensitive nucleus, allowing faster acquisition.

The anti-phase magnetisation on S can be used to record the chemical shift of the S spin, or it may be passed on to another building block for further manipulation. The reverse INEPT simply converts transverse S-magnetisation back to observable I magnetisation typically at the end of the pulse sequence for direct acquisition.

An additional advantage of the application of isotopically labelled

molecules is the additional number of structural extra parameters which may be measured by NMR spectroscopy. These include three-bond proton-carbon and carbon-carbon coupling constants, and heteronuclear NOEs (HOEs). In oligosaccharide NMR spectroscopy $^3J_{CH}$ values across glycosidic linkages, have been correlated with the glycosidic torsion angles (Mulloy *et.al.*, 1988), (Tvaroska *et. al.*, 1989) and a preliminary $^3J_{CC}$ Karplus curve for glycosidic linkages has also been developed (Milton *et al* 1997). ^1H-^{13}C HOEs have been reported for oligosaccharides (Poppe *et al.* 1992) and (Kiddle and Homans *unpublished data*). However, these studies are often complicated by the requirement of differential carbon-13 and deuterium labelling of different sugar residues in an oligosaccharide.

1.6.7 NMR Investigations of Ligand-Macromolecule Interactions

1.6.7.1 Tightly Bound Complexes

Tightly associating glycoconjugate-protein complexes are rather atypical, however, there are a few examples in the literature (Schengrund and Ringler. 1989; Low 1996). If an isotopically labelled ligand is available determining the conformation of a ligand which forms a tight association ($K_D > 10^{-7}$ M) with its target protein is potentially trivial as it is possible to directly observe the resonances of the bound ligand providing the protein:ligand stoicheiometry is 1:1 and the correlation time of the complex does not cause excessive line broadening. Isotope-edited NOESY spectra may be used to determine intra-ligand NOEs. Intermolecular NOE contacts are normally detectable in these complexes but it may be difficult to assign the specific protein resonances which form the contacts. Aromatic rings are often present in the vicinity of glycoconjugate binding

sites and large ring-current shifts (Δ 0.5-5 ppm) may be measured if the ligand is tightly bound (Low 1996).

1.6.7.2 Ligands in Fast to Intermediate Exchange

Most Lectins bind their oligosaccharide receptors with K_D values typically between ~10^{-3} to 10^{-6} M. Carbohydrates are thus rapidly associating and dissociating with the lectin many times during the lifetime of the NMR experiment. The resulting NMR spectrum is frequently a weighted average of the the free and bound states. The most obvious illustration of this the line broadening of resonances of a ligand experiencing exchange.

1.6.7.3 Binding-Induced Relaxation Enhancements and Differential Line Broadening Effects

In protein-ligand interactions, we usually deal with three species: the ligand, L, the protein, P, and the reversible molecular complex, LP (Figure 1.10). The binding equilibrium is characterised by the thermodynamic dissociation constant $K_D = k_{off}/k_{on}$, where k_{off} and k_{on} are the specific rate constant for the ligand dissociating and associating with the protein-ligand complex, respectively. Upon complex formation, a ligand nucleus often experiences a different magnetic environment depending on the location of the nucleus in relation to other nuclei in the bound ligand and in the bound protein. The environmental differences are reflected as differences in both the NMR T_2 relaxation times and the frequency separation for different ligand nuclei, thereby causing the observed differential line broadening effects on ligand resonances. The resonance line shapes for an exchanging system can be very complicated depending on the exchange kinetics and the concentrations of the exchanging

species.

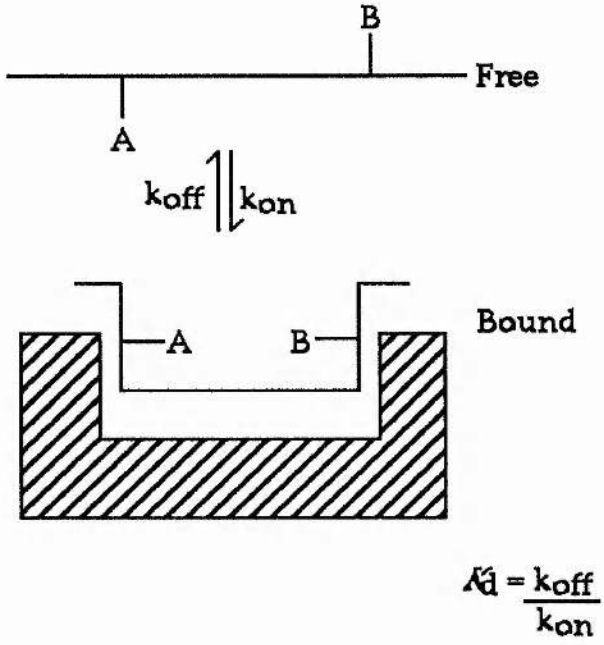


Figure 1.10: Schematic representation of a protein-ligand complex (LP) in dynamic exchange with a large excess of the free ligand (L). Protons in the free state may be too far apart for a measurable NOE to develop. However, protons A and B are close enough in the bound state for an NOE to be measured, however, measurements are made in the free state with magnetisation *transfers* achieved through combinations of chemical exchange and cross-relaxation mechanisms.

For an exchanging system where there is a large excess of free ligand present (or where the fraction of free ligand (α_f) is ~1) the resonance of each ligand is predominately a single signal (for an uncoupled nucleus) with a line width Δ is given by the Swift and Connick (1962) equation:

$$\pi\Delta = \frac{1}{T_{2,f}} + \alpha_b k_{off} \frac{\frac{1}{T_{2,b}}\left(\frac{1}{T_{2,b}} + k_{off}\right) + \delta\omega_b^2}{\left(\frac{1}{T_{2,b}} + k_{off}\right)^2 + \delta\omega_b^2} \qquad \textit{Equ 1.15}$$

where $T_{2,f}$ and $T_{2,b}$ are the NMR transverse relaxation times of the nuclei

in the free and bound states, respectively, α_b is the fraction bound, and $\delta\omega_b/2\pi$ is the frequency separation between the free and bound nuclei. Equation 1.15 indicates that exchange line broadening is mediated by the value of k_{off} and by the smaller $T_{2,b}$ for the bound ligand. However, more pronounced line broadening would be introduced by a larger difference between the free and the bound chemical shifts ($\delta\omega_b/2\pi$) regardless of the exchange kinetics. Thus, differential resonance broadening can been used to identify the contact sites on ligands that are interacting with the protein.

The influence of ring current shifts may also be extracted from the proton NMR spectra of oligosaccharides undergoing fast to intermediate exchange (Richardson *et al.* 1997). However, in contrast to the large changes in chemical shifts often observed for tightly bound ligands, the effects of ring current shifts in rapidly exchanging systems is more subtle and are often only evident by measuring ligand chemical shifts for a series of samples with increasing protein concentrations.

1.6.7.4 Transferred NOE Experiments

The transferred NOE or TRNOE is a particular example of the NOE in the presence of exchange of a ligand between free solution and a bound state in which it is complexed to a protein (Clore and Gronenborn 1982, 1983a) (and reviewed extensively in (Clore and Gronenborn 1983b; Campbell and Sykes 1993; Ni 1994; Ni and Scheraga 1994)). The aim of the TRNOE is to measure negative NOEs on the easily detectable free or averaged ligand resonances following the irradiation of other ligand resonances in order to obtain conformational information on the bound ligand (*Figure 1.10*).

As oligosaccharides are often in fast exchange with their receptor proteins, a variety of glycan-protein interactions have been probed by TRNOE experiments (Umenmoto *et al.* 1988; Bevilacqua *et al.* 1990; Bevilacqua *et al.* 1992; Bundle *et al.* 1994; Cooke *et al.* 1994; Weimer and Peters 1994; Andrews *et al.* 1995; Asensio *et al.* 1995; Milton and Homans 1995; Scheffler *et al.* 1995; and Richardson *et al.* 1997). The basis of the TRNOE is that σ^B, the cross-relaxation rate in the slowly tumbling bound state, is generally much larger than σ^F, the corresponding cross-relaxation rate constant in the rapidly tumbling free state. Thus, under conditions of fast exchange on the relaxation time scale, $<\sigma>$ is dominated by σ^B, so that the measured enhancements reflect the geometry of the ligand in its bound state. For fast exchange on the relaxation time scale, $<\sigma>$ is given by

$$<\sigma_{IS}> = N^F\sigma^F_{IS} + N^B\sigma^B_{IS} \qquad \textit{Eqn 1.16}$$

where N^F and N^B are the number of molecules in the free and bound states, respectively and σ^F_{IS} and σ^B_{IS} are cross relaxation rate between I and S in the free and bound states, respectively. The condition that $<\sigma>$ should be dominated by σ^B simply corresponds to the inequality:

$$|N^B\sigma^B_{IS}| \gg |N^F\sigma^F_{IS}| \qquad \textit{Eqn 1.17}$$

The extent to which this inequality is fulfilled depends on the relative correlation times (τ_c) of the free and the bound states, the distances r_{IS} in both states, and the equilibrium dissociation constant K_D. If the complex

is only weakly bound, then N^B will be small, and the TRNOE will disappear simply because there is insufficient ligand present in the bound state to contribute significantly to the overall relaxation. Clearly, the minimum value of K_D required for a viable TRNOE will vary depending on values of σ^B_{IS} and σ^F_{IS}. However, the more slowly tumbling the complex the larger $|\sigma^B_{IS}|$ becomes, consequently the simpler it is to satisfy Eqn 1.17. One advantage of this condition is that ligands in complex with very large proteins, such as the photon receptor rhodopsin (> 1000 kDa)(Ni 1994), which would normally be too slowly tumbling to study by conventional NMR spectroscopy may be probed by measuring TRNOEs.

At the opposite extreme, the TRNOE depends also of the *rate* of exchange between free and bound states being sufficient for an appreciable magnetisation flux to occur between them. If S^B is irradiated, the condition for appreciable enhancement transfer to occur from the enhanced bound signal I^B to the free signal I^F :

$$k_{on}[P] \gg R^F_I \qquad \textit{Eqn 1.18}$$

Where R^F_I is the relaxation rate of I in the free state. Similarly, if the signal S^F is irradiated selectively, the saturation transfer from S^F into signal S^B is also a requirement for a viable TRNOE, and here the relevant equality is

$$k_{off} \gg R^{B}_{S} \qquad \textit{Eqn 1.19}$$

where R^{B}_{S} is relaxation rate of S in the bound state. Inequalities Eqn 1.17 and Eqn 1.18 are likely to be broken for very tightly bound complexes. Thus for TRNOEs to be observed inequalities Eqn 1.17, 1.18 and for certain experiments Eqn 1.19 must be observed.

The buildup and decay of TRNOEs over a series of NOESY mixing times can also be used to study the relative mobilities of ligand residues when bound to a protein. For a typical example, TRNOEs which show an increase in intensity followed by decay over a small range of mixing times ($\tau_M \leq 100$ ms) indicates the presence of spin diffusion as a result of the restricted mobility of this residue and the close proximity of the protein. In comparison, TRNOEs which show a significant increase in intensity even up to relatively long mixing times ($\tau_M > 500$ ms) are characteristic of a group which has significant conformational mobility in the bound state. Thus, it is possible to map oligosaccharide residues which make significant contact with their receptor lectin (Casset *et al.* 1997).

Due to the restricted mobility of a ligand in the bound state and the close proximity of often very large proteins, spin-diffusion effects play a very important role in the TRNOE phenomena and failure to take spin diffusion effects into account has resulted in misinterpretation of TRNOESY data. A notable example of with was the work of Glaudermans and colleagues (1990) who examined the conformation of a disaccharide in complex with an antibody Fab fragment. The original study TRNOESY data suggested that there was a large conformational

change between the free and the bound states. However, in a following paper (Arepalli *et al.* 1995) the authors concluded that the inter-residue TRNOE which was indicative of this new bound state conformation was in fact an artifact of spin-diffusion and was a result of magnetisation transfer via the protein. Such indirect (TR)NOEs may be identified by measuring transferred NOE enhancements in the rotating frame (TRROE). In the rotating frame σ_{ROE} is positive regardless of the rate of rotational reorientation and gives rise to cross-peaks with negative phase. ROE-ROE relays (equivalent to spin diffusion in NOESY spectra) will always lead to crosspeaks having sign $(-1)^m$, where m is the number of ROE transfers (Farmer II *et al.* 1987), i.e. ROE transfer from protons A to C via a proton B is characterised by m=2 and yields a positive cross-peak.

1.7 Carbohydrate conformation

Carbohydrate conformations (reviewed in (Stoddart, 1971)) are defined by the conformations of individual residues and their linkage geometries.

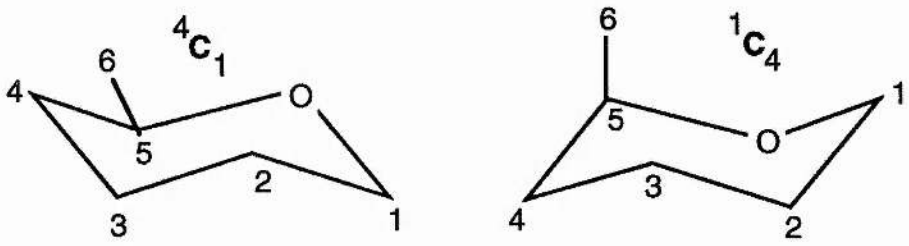


Figure 1.11: Chair conformations of pyranose rings, showing atom numbers.

Hexapyranose rings occur predominantly in the chair conformation. For the purposes of description, the atoms are numbered with the ring oxygen numbered zero and the carbon atoms follow sequentially from

the anomeric carbon, atom one. A reference plane is defined as passing through four of the non-hydrogen atoms in the ring, with the lowest numbered atom lying above or below the plane (IUPAC-IUB, 1980). Alternative conformations are then described by a letter describing the conformation ('C' for chair form). Preceding superscript and subscript numbers denote the atoms lying respectively above and below the plane of the ring (*See Figure 5.11*). The $^4C_1$ conformation predominates in D-pyranose saccharides, due to the relief of steric crowding. D-arabinopyranose in contrast adopts the $^1C_4$ conformation, due to the different distribution of hydroxyl groups, and lack of a 6-carboxy moiety (Rees, 1977). *N*-acetyl neuraminic acid adopts a similar conformation of D-arabinopyranose but is defined as $^2C_5$ because of the different numbering scheme used for this residue. Measurement of homonuclear spin-coupling constants does not indicate large deviations from theoretical predictions, hence pyranose rings are internally rigid structures.

Linkage geometry between rigid or partially flexible sugar residues is described by reference to the orientation about the glycosidic linkage. IUPAC-IUB (1970) describes this in terms of the heavy atoms about the linkage position, however in discussing NMR-derived conformations, it is more convenient to define them with respect to protons, as follows: ϕ_H=H1-C1-O1-Cx, ψ_H=C1-O1-Cx-Hx. The subscript '$_H$' distinguishes these definitions from the conventional IUPAC definitions, and 'Cx' and 'Hx' refer to aglyconic atoms. According to the Klyne-Prelog (1960) convention, an angle is defined as zero when the first and last bonds are

syn-periplanar. An angle is positive when a clockwise rotation of the atom at the rear of the field of view occurs.

1.7.1 Theoretical Determination of Conformation

The use of semi-quantitative NOE measurements in the determination of structure in proteins and nucleic acids is now well established. In contrast, 3D structure determination in oligosaccharides is limited due to the small number of inter-residue NOEs available, even with the inclusion of hydroxyl restraints. For this reason it is often necessary to apply theoretical analysis of computer-generated models. The atomic coordinates of an input structure are optimised to determine a minimum energy conformation. Methods for conformational searching and energy calculation procedures are reviewed in (French and Brady, 1989).

Full *ab initio* quantum mechanical computations of oligosaccharide conformations are unattainable, due to the computationally intensive calculations required, even for a small disaccharide. Hence computations of oligosaccharide conformation have relied almost entirely on molecular mechanical methods (Homans *et. al.*, (1987), Yan and Bush, (1990), Brady, (1987), Brady, (1986), Madsen *et. al.*, (1990)). This approach treats the molecule concerned as a number of point masses and charges, and a series of springs for bonds between them. Empirical data from X-ray crystallography or *ab initio* studies of small organic molecules provides optimum values for bond-lengths and torsion angles, deviation from which increases the energy potential of the model. Classical mechanics can then be used to derive an energy function for the model, allowing identification of lower energy configurations by differentiation. Additional factors affect the application of these methods to

carbohydrates, particularly the exo-anomeric effect (reviewed by Tvaroska and Bleha, 1989), the effect of which is to make the aglyconic carbon take up orientations of $\phi_H \sim +60°$ in β-D-glycosides, and $\phi_H \sim -60°$ in α-D-glycosides.

Early conformational calculations by Lemieux *et al* utilised the hard sphere exo-anomeric (HSEA) approach (Lemieux *et. al.*, 1980), (Thogerson *et. al.*, 1982). These calculations utilised a simple forcefield, including a Kityagordski term to describe nonbonded interactions, rigid ring geometries, neglect of partial atomic charges, and a torsional potential for the exo-anomeric effect. Results in good agreement with experimental and theoretical considerations were obtained for the blood-group oligosaccharides (Lemieux *et. al.*, 1980), (Thogerson *et. al.*, 1982), sucrose (Bock and Lemieux, 1982), and structures related to the complex-type *N*-linked glycans (Bock *et. al.*, 1982). These studies posit oligosaccharides as having essentially fixed orientations about the glycosidic linkages, with the exception of 1-6 linkages for which a number of orientations were predicted. Good agreement was obtained between these studies, and NMR or X-ray crystallographic experimental data. However in extending this approach to structures larger than 4-5 residues, agreement is poorer, and it becomes apparent that a degree of conformational flexibility may be present.

HSEA calculations give good results for static, minimum energy, structures. However, to investigate dynamic models, where the potential surface encompasses a wide range of ϕ and ψ angles, it now appears that a full molecular mechanical forcefield is more appropriate. The AMBER

parametrisation (Weiner *et. al.*, 1984), (Weiner *et. al.*, 1986), is a full forcefield, originally for nucleic acids and proteins, and has been extended to include carbohydrates by Homans (Homans, 1990), based upon work by Ha et. al. (1988a). A full molecular mechanical forcefield allows explicit inclusion of solvent water, in contrast to calculations done *in vacuo*.

HSEA calculations suggest deep potential wells, supporting the concept of rigid conformations. However, allowing all bonds, angles and torsion angles to vary during minimisation (flexible geometry) increases the area of conformational space accessible to a given glycosidic linkage (Scarsdale *et. al.*, 1988), (Homans *et. al.*, 1987), (French, 1989), (Ha *et. al.*, 1988b), (Homans *et. al.*, 1989), (Yan and Bush, 1990), (Imberty *et. al.*, 1989). Energy contour map calculations (Ramachandran plots) in which ϕ_H and ψ_H are systematically varied in small steps (10° or 15°) and the geometry optimised for this value of the glycosidic angles to produce relaxed potential surfaces, allows exploration of the extent of accessible conformational space. The results may be presented as an iso-energy contour map, in which occupation is temperature dependent, governed by the Boltzmann law (French and Brady, 1989). Calculation of relaxed maps is preferable to the use of a rigid-residue approach, since bad contacts may be relieved by relatively small changes in the internal coordinates (French, 1988), (Ha *et. al.*, 1988b), (Goebel *et. al.*, 1970). Clearly this approach becomes impractical in larger molecules, since the number of degrees of freedom rapidly increases both the number of calculations that must be carried out, and consequent computation time. In addition, the problem of false or multiple minima is accentuated, since it cannot be ensured that at each ϕ/ψ point the molecule is at a minimum energy.

An alternative method for the determination of accessible conformational space is the calculation of molecular dynamics simulations (Homans *et. al.*, 1987), (Yan and Bush, 1990), (Brady, 1987), (Brady, 1986), (Madsen *et. al.*, 1990). Studies on the behaviour of model saccharides during molecular dynamics simulations has highlighted several important points. Firstly, whilst structures may be restricted to low-energy regions of conformational space, they are by no means rigid. A number of conformations may exist independently during the period of a simulation, each consistent with experimental data. Secondly, internal motions occur on a timescale of picoseconds, substantially faster than the rate of overall tumbling for even moderately-sized molecules. The timescale of the build-up of the experimental NOE is of the order of milliseconds, whilst current computing technology renders the simulation of periods greater then tens of nanoseconds impractical. An additional advantage of MD simulations is that NMR parameters such as NOEs and spin-coupling constants may be back-calculated and compared with those measured experimentally. Significant improvements in the back-calculated NOEs and spin coupling constants have been observed when MD simulations are coupled with experimentally derived distance restraints. Distance constraints in restrained MD simulations take the form of flat bottomed harmonic restraint with a weak force constant (20 kJ mol^{-1}). The 'flat bottom' where no restraint is applied is there to reflect the inherent uncertainty of the NOE derived distances and because it is not possible *per se* to distinguish between a fixed internuclear distance and a variable internuclear distance caused by motional averaging. Distance restraints serve to restrict the motion of the MD simulation to areas of the potential energy surface which are well parametrised (i.e.

around global minima) and represent the minimum extent of torsional variation consistent with the experimental data. Unrestrained MD simulations, in comparison, represent the maximum torsional freedom consistent with the force field which often takes the molecules into regions of the potential energy surface divorced from the global minima which are poorly parametrised.

A novel strategy in the modelling of glycoconjugates is the application of 'time-averaged restraints' (Low, 1996). In this model, it is averaged rather than instantaneous structural parameters that are compared to experimental estimates. For practical reasons, the time average is taken over a memory time τ, so that an effective NMR distance might be (Torda, *et al.* 1990) given by:

$$r^{eff}(t)^{-6} = \frac{1}{t}\int_0^{\tau} e^{-t'/\tau}[r(t-t')]^{-6}\,dt' \qquad \textit{Eqn 1.20}$$

Where r *and* r^{eff} are the distance and averaged distance, respectively, and t is time. Since only the average (and not the instantaneous) distances need to satisfy the constraints, the restrictions on the conformational space sampled are generally less severe than for "conventional" restrained MD simulations. This becomes particularly important as the number of experimental constraints detected by NMR spectroscopy for oligosaccharides increases following developments in the measurement of NOEs to exchangeable protons and heteronuclear NOEs. This increase in distance data will likely produce sets of conflicting distance constraints which can only be modelled well by time-averaged restraints.

Determination of the minimum energy structure may be reached by means of restrained simulated annealing (Clore *et. al.*, 1985). Convergence upon similar final conformations, from a number of different starting conformations ensures that a larger proportion of conformational space has been sampled compared to minimisation by differentiation of the energy function.

1.8 Outline of investigation

In this investigation, the solution conformation(s) and dynamic properties of derivatives of the glycolipid receptors for the bacterial toxins, cholera/heat labile and pig edema toxin are determined. To aid future investigations of biological glycoconjugates, carbon-13 enriched oligosaccharide moieties commonly found in glycolipids and glycoproteins were synthesised and multidimensional heteronuclear NMR studies were performed on test systems.

In chapter 2 we investigate the three dimensional structure of the oligosaccharide moiety of the ganglioside G_{M1} by measuring NOE intensities and comparing them against values calculated from computer models.

Chapter 3 studies the structure and dynamics of a chemically synthesised derivative of globotetraosylceramide, Gb_4-OEtTMS. Conventional inter-residue NOEs measured between C-linked protons were complimented by additional NOEs measured between exchangeable and nonexchangeable protons in aqueous solution.

Chapter 4 details the chemoenzymatic synthesis of carbon-13 enriched sialyltrisaccharides α2-3 sialyllactose (oligo-G_{M3}: Neu5Acα2-3Galβ1-4Glc) and α2-3 *N*-acetyl sialyllactosamine (Neu5Acα2-3Galβ1-4GlcNAc).

In Chapter 5 high resolution multidimensional heteronuclear NMR experiments are used to determine the 3D structure and dynamics of ^{13}C enriched oligo-G_{M3} in free solution and in complex with the plant lectin wheat germ agglutinin.

Chapter 6 describes a preliminary study of the conformation and dynamics of the glycan moiety of an intact glycoprotein, the Fc fragment of immunoglobulin G by incorporating *N*-acetyl [U-^{13}C] neuraminic acid α2-6 [U-^{13}C] galactose into the terminal position of the complex glycan.

Chapter 2

Solution dynamics and Structure of the Oligosaccharide Moiety of Ganglioside G_{M1}

Abstract

The solution dynamics of the oligosaccharide moiety of ganglioside G_{M1} have been determined by use of a combination of ^1H rotating frame Overhauser effect measurements and restrained molecular dynamics simulations. It is found that the Galβ1-3 and Neu5NAc moieties which are primarily recognised by cholera toxin both exhibit considerable torsional flexibility about their respective glycosidic linkages. A comparison with the bound state conformation of the ganglioside in association with cholera toxin B-pentamer, shows that a low energy conformation of the oligosaccharide, which closely approximates the global minimum, is selected upon binding.

2.1 Introduction

Cholera toxin, an 85,000-dalton protein, is composed of two subunits, A and B (for review, see (Spangler, 1992) and (Hol *et al.* 1995)). The A subunit contains two polypeptide chains, A1 and A2 linked by a disulphide bond. A1 has enzymatic activity which catalyses transfer of an ADP-ribose moiety from NAD to the $G_{s\alpha}$ regulatory component of adenylate cyclase in susceptible cells, resulting in constitutive production of adenylate cyclase. The B subunit consists of a noncovalent pentamer of identical polypeptide chains, which binds to the oligosaccharide moiety of cell surface monosialoganglioside G_{M1} (oligo-G_{M1}, Galβ1-3GalNAcβ1-4(Neu5Acα2-3)Galβ1-4Glc) (*See Figure 2.1*) (van Heynigen *et. al.*, 1971). Recently, the crystal structure of the cholera toxin B-pentamer - G_{M1}-oligosaccharide complex has been reported (Merritt *et. al.* 1994). The majority of interactions between the receptor and the toxin involve the 2 terminal sugars of G_{M1}, galactose and *N*-acetyl neuraminic acid, with a smaller contribution from the *N*-acetylgalactosamine residue. Since oligo-G_{M1}, in common with most oligosaccharides, might be expected to exhibit significant conformational mobility in solution, it is of interest to determine which of the various solution conformations is recognised by the toxin. The solution structure of G_{M1} and its oligosaccharide have been the subject of previous investigations. Early theoretical studies (Sabesan *et. al.* 1984) using NMR and hard sphere exoanomeric effect (HSEA) calculations led to the prediction of the preferred conformation of oligo-G_{M1} in solution, but did not consider explicitly the dynamic interconversion between various alternative conformers. Conversely, in a more recent study the dynamics of the oligosaccharide headgroup of

G_{M1} was considered in more detail (Acquotti*et. al.*, 1990), but the solvent used in that study (dimethylsulphoxide), precludes a straightforward study with the oligosaccharide conformation in the crystal complex. We therefore sought to examine the solution structure and dynamics of oligo-G_{M1} using a combination of high-resolution NMR together with molecular dynamics simulations.

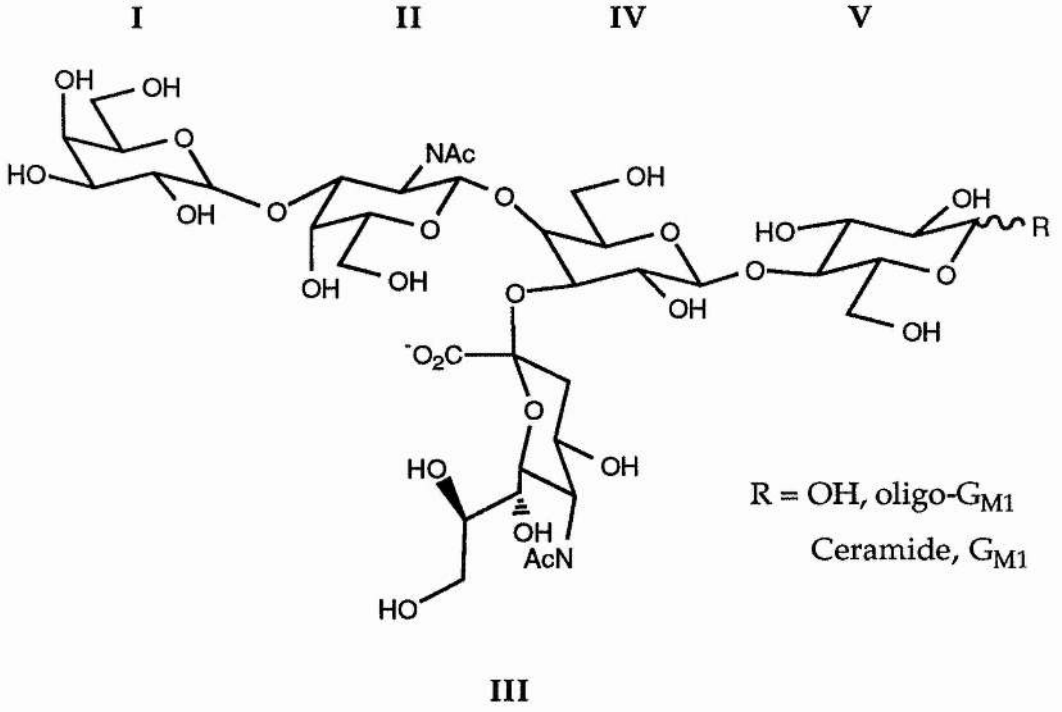


Figure 2.1: Illustration of the structure of oligo-G_{M1} including residue numbering system: I-Gal, II-GalNAc, III-Neu5Ac, IV-Gal, and V-Glc.

2.2 Materials and Methods

Sample Preparation

oligo-G_{M1} was prepared by digestion for 48 hrs. of 5 mg. ganglioside G_{M1} with 10 mU ceramide glycanase (Sigma) in 10 mM sodium acetate buffer, pH 5.5, containing 0.4% (w/v) of Triton X-100. The glycan was purified by

extraction of detergent into toluene, followed by extraction of lipid by chloroform/methanol (2/1, v/v) partition. The aqueous phase was flash evaporated three times from 99.96% Deuterium Oxide, followed by dissolution in 700 ml 99.96% Deuterium Oxide for NMR studies.

NMR Experiments

Homonuclear ^{1}H-^{1}H offset-compensated ROESY experiments were performed as described previously (Rutherford *et. al.* 1994), using a spin-lock field of 2.5 kHz. which was applied for 300 ms. The carrier was located at the downfield end of the spectrum during the spin-lock period to minimize coherence transfer effects.

Molecular Modelling

Computer models were assembled using Biosym's Insight II molecular graphics package (v 2.8, Biosym Technologies, San Diego, CA) running on a Silicon Graphics Indigo 2 Challenge computer. Random structures were generated by dynamical quenching. An initial structure was built with pyranose rings in the $^{4}C_{1}$ ($^{2}C_{5}$ for Neu5Ac) chair conformation with trial values of phi (ϕ) and psi (ψ), and subjected to 200 ps of unrestrained molecular dynamics at 750K, during which the torsional terms are scaled by a factor of 7 to prevent excessive ring puckering. A random structure was saved every 10 ps. Energy minimisation by restrained simulated annealing was achieved as follows: models were equilibrated for 10 ps with a thermal bath at temperatures 500K, 450K, 350K, 300K, and then successively for 1 ps in decreasing steps of 10K, followed by a further 1 ps at 5K. The system was minimised using the steepest descents algorithm until the maximum derivative was less than 0.04 kJ $mol^{-1}Å^{-1}$. Restraints

listed in table 4 were applied as a biharmonic function. NOE contacts were arbitrarily assigned as strong (1.8Å - 2.7Å), medium (1.8Å - 3.3Å) or weak (1.8Å - 5.0Å). TRNOE contacts were assigned as strong (1.8Å - 2.7Å) or weak (1.8Å - 5.0 Å).

All dynamics simulations were performed *in vacuo* with a dielectric constant (ε) of 80.0. Restrained molecular dynamics simulations lasted a total of 510 ps real time, with the last 500 ps used in further analysis.

The torsional angles ϕ and ψ are analogous to ϕH, ψH in IUPAC convention, and are defined as H1-C1-O1-CX, C1-O1-CX-HX , where CX and HX are aglyconic atoms. Torsion angles for the glycerol moiety of Neu5Ac are defined as ϕ_1 = H6-C6-C7-H7 and ϕ_2 = H7-C7-C8-H8.

2.3 Results and Discussion

Solution Structure of Oligo-G_{M1}

Proton assignments were made by Dr. J. M. Richardson and were identical to those in the literature (Acquotti, 1990). Initially, proton-proton through-space distance constraints were mapped in oligo-G_{M1} by use of rotating frame Overhauser effect spectroscopy. A total of eight inter-residue ROEs were obtained. Of these, one was an indirect ROE which was not used as a constraint, and the remaining seven were quantified empirically as 'weak', 'medium' and 'strong' for the purpose of restrained MD simulations below. In addition, all eight ROEs were quantified by integration of crosspeak volumes for later comparison with theoretical averaged ROE values (*vide supra*). A series of ten pseudorandom

geometries of oligo-G_{M1} were then computed by use of a dynamic quenching procedure, and each geometry was subjected to dynamical simulated annealing, with inclusion of the seven ROE constraints.

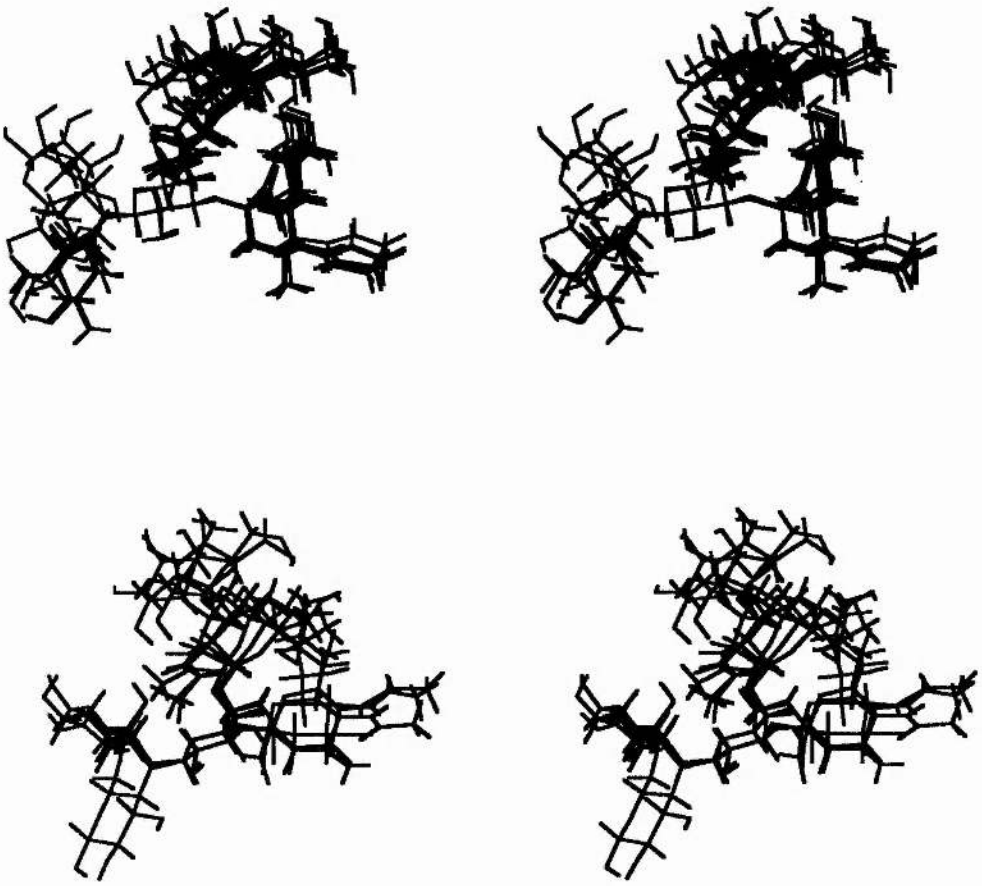


Figure 2.2: Energy minimised structures derived from restrained simulated annealing calculations. Shown are the annealed coordinates from ten initial random structures, which formed two distinct families differing primarily in the orientation about the Neu5Acα2-3Gal linkage; (above) ϕ,ψ ~ -165°,-30° and (below) ϕ,ψ ~ +140°, -20°..

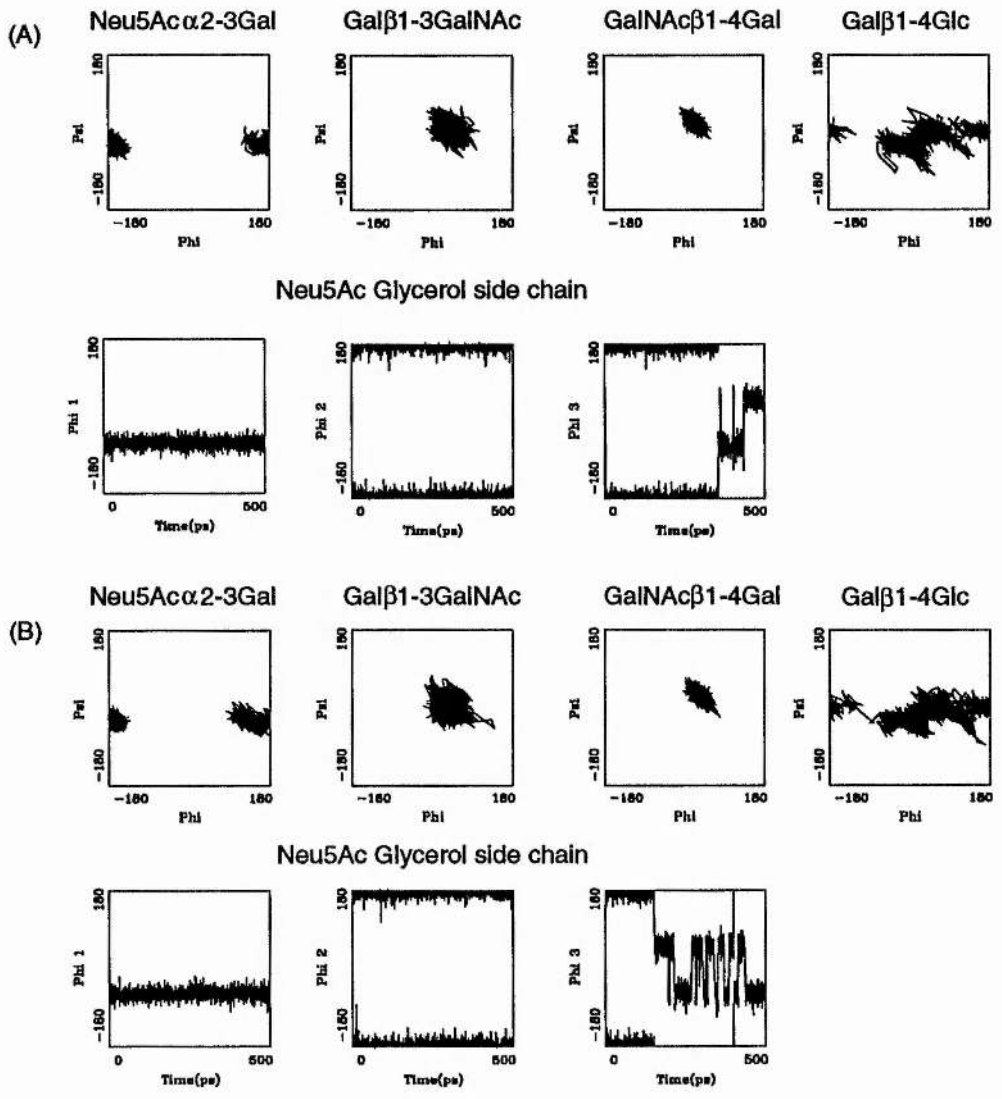


Figure 2.3: Time-dependent variations in torsion angles over the timecourse of the 500 ps restrained MD simulations of oligo-G_{M1}. Panels at the top refer to the simulation starting with the lowest energy structure from the 'global minimum' family of structures (at top in *Figure* 2.2), and panels at the bottom refer to a similar simulation starting with the lowest energy structure from the 'local minimum' family of structures (at bottom in *Figure* 2.2).

above as biharmonic distance restraints. This procedure resulted in two distinct families of structures (*Figure* 2.2). The first, comprising seven structures, was found to have the lowest energy and was thus defined as the 'global minimum' family. The second family, which was less well defined, comprised three structures whose average energy was 8.4 kJ/mol higher than the first. This family differed from the first primarily in a substantially different orientation for the Neu5Ac unit, and was defined as the 'local minimum' family. There were no constraint violations in any of the annealed structures.

In order to determine the extent of torsional fluctuation about the global and local minimum structural families above, the lowest energy structure from each of the two families was arbitrarily chosen and used as input to a 510 ps restrained molecular dynamics simulation in vacuo, using the same constraints as above. As discussed at length previously (Rutherford, T. J. and Homans, S. W. 1994), the resulting fluctuations about ϕ and ψ for each glycosidic linkage (*Figure 2.3*) represent the minimum extent of torsional variation consistent with available experimental data. Each of the restrained simulations suggests that the Galβ1-4Glc and Galβ1-3GalNAc linkages, in particular, exhibit significant torsional flexibility about their respective glycosidic torsional angles ϕ and ψ. Conversely, more limited torsional oscillations exist about the GalNAcβ1-4Gal linkage. There appear to be two discrete minima about the Neu5Acα2-3Gal linkage, and both are populated on the 500 ps timescale, irrespective of whether the simulation is initiated from $\phi,\psi \sim -165°, -30°$, which corresponds to the minimum energy conformation in the global

minimum family of structures, or $\phi,\psi \sim +140°, -20°$, corresponding to the minimum energy conformation in the local minimum family of structures. This contrasts with a previous study in dimethylsulphoxide (Acquotti*et. al.* 1990), where only a single conformation about the Neu5Acα2-3 linkage was predicted.

In order to determine whether the torsional fluctuations in figure 2.2 are consistent with experimental data, time-averaged ROEs were back-calculated over the timecourse of the MD simulations using a formalism appropriate for internal motions which are fast with respect to the overall tumbling time of the molecule (Homans, S. W. and Forster, M. J. 1992).

Table 2.1. Comparison of experimental ROEs for oligo-G_{M1} with time-averaged values derived from MD simulations.

Galβ1-3GalNAcβ1-4(Neu5Acα2-3)Galβ1-4Glc
I II III IV V

ROE constraint		Experimental ROE(%)[1]	Theoretical ROE(%)[2] Simulation A	Simulation B
I-H1 - II-H3	(S)[3]	7.1*	4.9	6.3
I-H1 - II-H2	(W)	0.8	0.2	0.2
I-H1 - II-CH3	(W)	0.3	0.3	0.3
II-H1 - IV-H4	(S)	5.9	6.1	4.2
II-H1 - III-H8	(M)	1.5*	1.3	3.9
II-CH3 - IV-H2	(M)	1.1	2.8	2.8
IV-H3 - III-H3ax.	(S)	4.1	7.1	4.1
IV-H3 - III-H3eq.		-1.2	-1.2	1.1

[1]Intensities normalised to theoretical ROE between **III**-$H3_{ax}$. and **III**-$H3_{eq}$.

[2]Time averaged values for MD simulation starting with global minimum family (simulation A) and local minimum family (simulation B) of structures, respectively, computed using an isotropic rotational correlation time of 0.35 ns.

[3]Qualitative magnitude of ROE used as restraint: (S): Strong; (M): Medium; (W): Weak. See text for further details.

*Accurate quantification precluded by resonance overlap

In contrast to previous calculations where a spin-pair approximation was applied in computing theoretical ROEs (Rutherford. T. J. *et. al.* 1994), in the present study we performed a full relaxation matrix calculation on the entire proton spin system of the molecule. This approach was clearly mandatory in view of the fact that one of the measured ROEs in oligo-G_{M1} was of opposite sign than that expected for a 'direct' ROE, indicating an indirect 'three-spin effect'. Such effects can only be computed adequately via a full relaxation matrix approach.

Time-averaged ROEs were computed independently from the MD simulations for each family of structures. A single overall isotropic correlation time for the molecule was assumed, and was obtained by fitting the ratio of theoretical diagonal peak to crosspeak (intra-residue ROE) intensities of the **III** $H3_{ax}$ and $H3_{equ}$ protons to the experimentally measured values. The experimental ROE intensities (i.e. crosspeak volumes) were then normalised to the theoretical ROE intensities to generate the absolute magnitudes of the experimental ROEs, and the resulting values are shown in table 2.1. These values did not differ substantially irrespective of whether an alternative intra-residue ROE (such as that between **I** H1 and H5) was used for calibration. Since conformations representative of both structural families are expected on the timescale of measurement of the ROE, the predicted experimental values will be some average of the theoretical values obtained from each simulation. As noted above, the primary difference between the two simulations lies in the conformation about the Neu5Acα2-3Gal glycosidic linkage, and since the rate of torsional variation is not fast with respect to the overall tumbling time of the molecule (0.35 ns), then the predicted

values of the ROEs become a weighted mean of the ROEs in each of the two conformational states. As discussed previously, agreement within a factor of two between experimental and theoretical ROEs can be considered satisfactory given the sixth power dependence of the ROE upon distance from the MD simulations. It can be seen by inspection of Table 2.1 that only the weak ROE between **I**-H1 and **II**-H2 differs substantially between the theoretical simulation and the experimental measurement. It is very probable that the experimental value is overestimated due to the difficulty in obtaining a volume integral for such a small crosspeak in the presence of noise. In general therefore the dynamic behaviour during either simulation is consistent with the experimentally measured direct ROEs. In contrast, there is a difference in sign between the ROE from **IV**H3 and **III**-$H3_{eq}$, and only the first simulation generates an indirect ROE which corresponds in sign to the experimental value. This ROE therefore provides a crucial distinction between possible models of solution behaviour of oligo-G_{M1}, and strongly suggests that the solution behaviour predicted in simulation A represents the dominant conformational state, i.e. with the conformation about the Neu5Aca2-3Gal glycosidic linkage at $\phi,\psi \sim -165°,-30°$. Assuming an error in experimental measurement of ROE values of 20%, the predicted population of the alternative conformation ($\phi,\psi \sim +140°, -20°$) for this linkage is no greater than 10%.

Oligo-G_{M1} exhibits limited torsional flexibility in aqueous solution. The GalNAcβ1-4Gal linkage is particularly highly constrained, primarily by the proximity of the Neu5Ac residue. While the Galβ1-3GalNAc linkage exhibits considerably greater conformational freedom, it nevertheless

appears to exhibit torsional oscillations within a single energy well. In contrast, the Neu5Acα2-3Gal linkage exists in two discrete conformational states. This appears to be the only major difference from the behaviour of the glycan in intact G_{M1} in dimethylsulphoxide solution (Acquotti *et. al.*, 1990), where only one conformational state was predicted. However, the conformation predicted in dimethylsulphoxide is the dominant conformer in aqueous solution.

The values of the glycosidic torsion angles of oligo-G_{M1} in association with the B-subunit of cholera toxin (Merritt *et. al.*, 1994) are compared in Table 2.2 with the average values of the glycosidic torsion angles from the solution structure in the present study. It is clear that the bound-state conformation of the glycan maps onto low energy regions of the ϕ,ψ potential surface for each linkage. In addition, the glycerol moiety of Neu5Ac, which is also involved in binding, appears to be bound in a conformation which is essentially identical to that observed in solution. As is apparent from figure 2.3, the torsional angles ϕ_1 and ϕ_2 for the glycerol moiety essentially adopt single values. In previous work on G_{M1} in DMSO, it was suggested that the glycerol moiety was stabilised by a hydrogen bond between III-C8-OH and either the carboxylic or ring oxygen, and between III-C7-OH and the *N*-acetamido carbonyl oxygen (Acquotti, D. *et. al.*, 1990). It might be anticipated that such hydrogen bonds would not be long-lived in solvent water. However, it has recently been demonstrated by Poppe and van Halbeek (1991) that the III-C8-OH hydrogen bond does indeed exist in aqueous solution. In the present study, we find that the values of the $^3J_{H,H}$ coupling constants for the glycerol sidechain of Neu5Ac in oligo-G_{M1} in aqueous solution are

essentially identical to those reported previously for G_{M1} in DMSO, and hence our data are consistent with previous work. In the crystal structure, the III-C7-OH hydrogen bond to the *N*-acetamido carbonyl is preserved, but however the III-C8-OH is hydrogen bonded via a water molecule to Gly 33 in the binding site.

It is remarkable that the solution conformation of oligo-G_{M1} is strongly preserved in the crystal structure of the complex. In those carbohydrate-protein complexes for which high-resolution data are available (Bourne *et. al.*, 1990a,b; Cygler *et. al*, 1991; Imberty *et. al.*, 1991 and Low, 1996), the carbohydrate is invariably perturbed from the lowest energy structure observed in solution, and hence the oligo-G_{M1} - Cholera toxin complex represents an unusual case. The relative rigidity of oligo-G_{M1}, together with the lack of conformational strain on binding, probably accounts for the very high binding affinity of cholera toxin for oligo-G_{M1}, since binding energy is not lost via conformational strain, nor in overcoming the configurational entropy which would be present in a more flexible ligand. Therefore, with regard to the rational design of inhibitors of the oligo-G_{M1} - cholera toxin complex, the present study predicts that there might be little gain in binding affinity by design of a oligo-G_{M1} analogue with reduced configurational entropy.

Table 2.2. Comparison of the average solution conformation and bound state conformation of oligo-G_{M1} in association with cholera toxin B-subunit.

Linkage	Crystal (deg.)	Solution(deg.)[a]
Galβ1-3GalNAc	57.7[b]	39.9 (17.4)[d]
	3.0[c]	11.6 (16.5)
GalNAcβ1-4Gal	48.2	23.8 (9.8)
	6.8	24.4 (9.8)
Galβ1-4Glc	-48.7	28.8 (68.3)
	-6.7	-12.5 (20.5)
Neu5Acα2-3Gal	-168.5	-130.0 (99.0)
	-30.8	-25.0 (10.2)
Neu5Ac ϕ1	-54.2	-60.6 (10.55)
ϕ2	-178.3	-169.9 (10.88)

[a]Values reported for dominant conformer about the Neu5Acα2-3Gal glycosidic linkage

[b] ϕ

[c] ψ

[d] Values in parentheses are the rms deviations for each torsion angle over the timecourse of the simulation.

2.4 Conclusions

Conventional two dimensional homonuclear rotating frame NOEs were used to model the minimum energy structures and dynamics of the oligosaccharide moiety of G_{M1}. The Neu5Acα2-3Gal linkage appears to oscillate between two minima in aqueous solution, as compared to the single solution conformer proposed in DMSO (Acquotti*et al* 1990). However, the dominant solution conformation (ϕ, ψ -165°, -30°) is preserved in both solvents. The observed 'indirect' ROE coupled with the rigorous evaluation of dynamics simulations, only possible with full

relaxation matrix calculations, allowed the discrimination between the relative populations of the global and local minimums.

Several drugs have been proposed for the treatment of cholera, travellers diarrhoea ((Spangler, 1992; Hol *et al*. 1995) and references therein), and epilepsy (McAuliffe and Hindsgaul, 1997) based on the G_{M1} structure. In the context of drug design, it is interesting to note the similarity between the dominant (global minimum) solution conformation and the bound conformation of oligo-G_{M1} in complex with the cholera toxin B-subunit (Merrit, 1994). It would appear that any anti-cholera therapeutics will have to mimic both the observed solution bound conformation of G_{M1} and very importantly, the constrained solution dynamics of oligo-G_{M1}. An important point must be taken into account when considering the the solid-state bound conformation of G_{M1}: in four out of the five binding sites cholera toxin posses, the electron-density for the carbohydrate was very disordered and only a fortuitous crystal contact stabilised the conformation of oligo-G_{M1} in a single binding site. This crystal packing interaction may perturb the conformation of G_{M1} and the solution bound conformation of G_{M1} remains to be determined but such studies would require the synthesis of a uniformly carbon-13 enriched ligand. The synthesis of moieties required for G_{M1} is discussed in chapter 4.

Chapter 3

Three-Dimensional Structure and Dynamics of the Oligosaccharide Moiety of an Analogue of the Globotetraosylceramide Gb_4

Abstract

The three dimensional structure and dynamics of the oligosaccharide moiety of a synthetic analogue (globotetraosyl *O*-ethyltrimethylsilane Gb_4-OEtTMS, (GalNAcβ1-3Galα1-4Galβ1-4GlcβO$(CH_2)_2$Si$(CH_3)_3$)) of the globotetraosylceramide Gb_4 (globoside, GalNAcβ1-3Galα1-4Galβ1-4GlcCer) was investigated in D_2O and in H_2O containing 15% acetone-d_6 by means of laboratory and rotating frame NOE experiments coupled with restrained molecular dynamics simulations. A previously unreported solution conformer about the GalNAcβ1-3Gal terminus was predicted using distance restraints derived from ROE contacts between C-linked protons in D_2O. A similar conformation was also proposed on the basis of NOE contacts between exchangeable and non-exchangeable protons. NMR data are consistent with a conformationally flexible molecule in free solution. The dominant region of conformational space sampled by the molecule coincides the predicted conformation of the sugar in complex with pig edema toxin.

3.1 Introduction

Globotetraosylceramide Gb_4 (GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-1Cer, *See Figure 3.1*) belongs to the globo series of glycosphingolipids found in a variety of mammalian tissues.

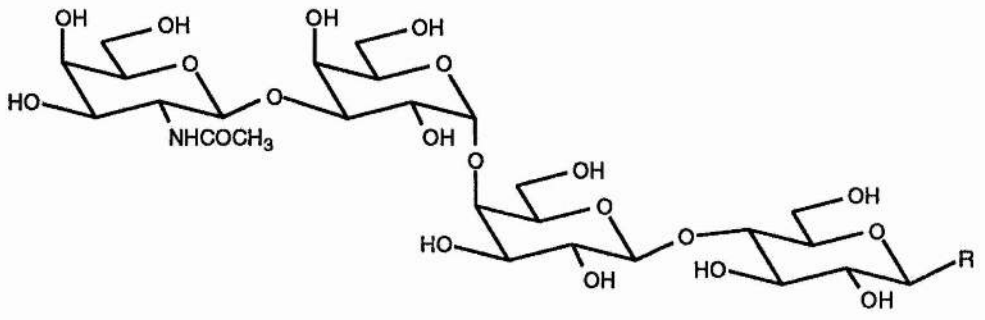


Figure 3.1: Schematic structure of the oligosaccharide moiety of Gb_4.

Globo-series glycolipids have been shown to play a role in many disease states. In particular Gb_4 and the truncated version globotriosylceramide Gb_3 (Galα1-4β1-4Glcβ1-1Cer) act as cellular receptors for a number of enteropathogenic bacterial toxins. *S. dysenteriae* shiga toxin and *E. coli* shiga-like toxins (verotoxins, VT) both have been shown to selectively bind Gb_3 (Jacewicz *et al.* 1986), whereas, the pig edema toxin (VT2e) complexes with Gb_4 and to a lesser extent with Gb_3 (DeGrandis *et al,* 1989). Interest in these toxins has recently been renewed following devastating out-breaks of food poising in the United States of America and Scotland. Contamination usually follows the consumption of poorly cooked beef. The disease is usually fatal in the very young or the elderly and often

results in heamolytic uraemic syndrome; the major cause of acute paediatric renal failure. As cellular adhesion of these toxins is a prerequisite for cellular invasion and disease, many research groups are currently investigating the possibility of synthesising small molecules which may inhibit cellular attachment (McAuliffe and Hindsgaul, 1997; Rougvie, 1997). A starting point for such studies has been to modify the natural carbohydrate receptors. Knowledge of the conformation that the sugar presents to the toxin prior to complex formation and of the structure of the complexed form are essential if a rational approach towards drug design is to be undertaken. An interesting observation from preliminary studies is that the removal of the ceramide tail results in a marked decrease in binding affinity ($K_D \sim 10^{-9}$ M and $\sim 10^{-3}$ M for the intact globoside and the free oligosaccharide moiety, respectively). Verotoxin-receptor interactions have been shown to be multivalent (Lindberg *et al* 1990), with binding affinities increasing depending on the number of oligosaccharide moieties present on a molecular surface. It has also been postulated that the ceramide tail may stabilise the conformation of the glycan moiety or interact with the protein directly thus aiding complex formation (Lingwood, 1993).

The conformation of an analogue of Gb_3 (Gb_3-OEtTMS, Galα1-4Galβ1-4GlcβO$(CH_2)_2Si(CH_3)_3$) both in free solution and in complex with verotoxin-1 (VT-1) was the subject of an investigation presented elsewhere (Richardson *et al* 1997). This current investigation centres on probing the solution conformation of Gb_4-OEtTMS which binds to pig edema toxin VT2e less strongly than intact Gb_4 itself (*vide supra*). Although there is little evidence for a role for Gb_4 in disease in man, a

derivative of Gb_4, amino-Gb_4 ($GalNH_2\beta1$-3Gal$\alpha1$-4Gal$\beta1$-4Glc$\beta1$-1Cer) is a proposed inhibitor of all pathogenic verotoxins (Nyholm *et al* 1996a). The solution behaviour of Gb_4-OEtTMS was investigated using ROESY data to generate distance restraints for molecular modelling. Additional restraints between exchangeable and C-linked protons were generated by acquiring NOESY spectra in H_2O at low temperatures (258K).

3.2 Materials and Methods

Sample Preparation

Chemically synthesised Gb_4-OEtTMS was provided by Dr. G. Magnusson. Samples studied in D_2O were prepared by dissolution/flash evaporation of Gb_4-OEtTMS (4 mg) three times from 99.96% D_2O (Aldrich), followed by dissolution in 500 µL 99.96% D_2O. For the observation of exchangeable protons, Gb_4-OEtTMS (2.5 mg) was dissolved in 700 µL distilled H_2O containing 15% 99.95% acetone-d_6 (Aldrich). The pH was adjusted to 6.5 by careful step-wise addition of 20 mM HCl and transferred to a 5 mm NMR tube (Wilmad 528pp). The sample was degassed by sonicating the NMR tube for approximately 1 minute.

NMR Experiments

NMR spectra were obtained at a ^1H reference frequency of 500.3 MHz with a Bruker AMX 500 or a Varian Unity$^+$ spectrometer equipped with a self-shielded z gradient triple resonance probe. Chemical shifts were referenced to TSP indirectly by setting the ^1H signal of the residual HOD or acetone-d_6 to 4.75 ppm or 2.25 ppm respectively.

The following spectra were acquired in D_2O with a probe temperature of 300K. A homonuclear ^1H two dimensional (2D) coupling correlated spectroscopy experiment (COSY) was recorded in magnitude mode with a spectral width of 2.2 kHz, 512 t_1 increments, 8 scans per increment with 2K complex points in t_2. The remaining spectra were recorded in the phase-sensitive mode with use of the States' (1982) method for quadrature detection in t_1. A 2D total correlation spectroscopy ((TOCSY) (Brüschweiler *et al.* 1988), HOHAHA (Subramanian *et al.* 1987) experiment was acquired with a spectral width of 2.2 kHz, a spin-lock time of 119 ms, 1024 t_1 increments, 8 scans per increment with 2K complex points in t_2. A 2D triple-quantum filtered COSY (TQF-COSY) experiment was acquired with 512 t_1 increments, 32 scans per increment with 2K complex points in t_2 and a spectral width of 2 kHz. Homonuclear 2D offset compensated rotating-frame nuclear Overhauser effect (ROESY) experiments were recorded as described previously (Rutherford *et al.* 1994). The rf carrier frequency was placed 500 Hz up field/down field relative to the highest/lowest field resonance to minimise coherence transfer effects. The effective field for spin-locking was 4 kHz and was applied for a mixing time of 300 ms. A total of 1024 t_1 increments, 32 scans per increment with 4K data points were accumulated with a spectral width of 3.7 kHz. A ^{13}C-^1H-correlated heteronuclear multiple-quantum coherence (HMQC) experiment was acquired with 512 t_1 increments, 64 scans per increment with 4K complex points in t_2. ^1H-^{13}C couplings evolving in t_2 were removed by GARP decoupling.

The following spectra were recorded in H_2O containing 15% acetone-d_6 at 258K. Excitation sculpted (ES) (Hwang and Shaka, 1995) 2D TOCSY,

NOESY, and ROESY experiments were recorded with a spectral width of 5.6 kHz, 512 t_1 increments, 32 scans per increment with 4K complex points in t_2. ES-NOESY and ES-ROESY experiments were acquired with a mixing time of 100 ms, a 4 kHz spin-lock field was applied during the ROESY mixing time. Hydroxyl exchange rates were measured using the method described by Adams, B. and Lerner, A. (1991), incorporating the water suppression method described in (Hwang and Shaka, 1995). Spectra were acquired with 16 scans, 16K complex data points, a sweep-width of 5.6 kHz and were recorded with mixing times varying from 0 to 0.1s. The relaxation delay was 10s. Line heights were measured using the kinetic analysis package included in the Varian VNMR software. Non-exchangeable proton spin-coupling constants were measured from an ES-1D ^{1}H spectrum acquired with a proton spectral width of 5.6 kHz and 30K complex points.

COSY data were multiplied in each dimension by unshifted sine-bell functions. All 1D and 2D time domain data were zero filled once, multiplied by a squared cosine bell function prior to Fourier transformation. NOESY and ROESY peak volumes were measured using the 2D integration routine contained in Varian software.

Conventions

The torsion angles ϕ(phi), ψ(psi) are analogous to ϕ_H and ψ_H in IUPAC convention, and are defined as H1-C1-O1-CX and C1-O1-CX-HX, where CX and HX are aglyconic atoms.

Molecular Modelling

Computer models were assembled using Biosym's Insight II molecular graphics package (v 2.8, Biosym Technologies, San Diego, CA) running on a Silicon Graphics Indigo 2 Challenge computer. Random structures were generated by dynamical quenching. An initial structure was built with pyranose rings in the $^4C_1$ chair conformation and with trial values of phi (ϕ) and psi (ψ), and subjected to 200 ps of unrestrained molecular dynamics at 750K, during which the torsional terms are scaled by a factor of 7 to prevent excessive ring puckering. An additional torsional restraint (412 $kJmol^{-1}rad^{-2}$) was required to hold the NH of the *N*-acetyl moiety in the *E* conformer. A random structure was saved every 10 ps. Energy minimisation by restrained simulated annealing was achieved as follows: models were equilibrated for 10 ps with a thermal bath at temperatures 500K, 450K, 350K, 300K, and then successively for 1 ps in decreasing steps of 10K, followed by a further 1 ps at 5K. The system was minimised using the steepest descents algorithm until the maximum derivative was less than 0.04 kJ $mol^{-1}Å^{-1}$. Restraints are listed in table 3.4. NOE contacts were arbitrarily assigned as strong (1.8Å - 2.7Å), medium (1.8Å - 3.3Å) or weak (1.8Å - 5.0Å).

Molecular dynamics simulations were performed *in vacuo* with a dielectric constant (ε) of 80.0. The AMBER force-field (Singh *et al.*, 1986) with a carbohydrate parameter set developed by Homans (Homans 1990) was used with the exo-anomeric potentials set to zero. Simulations lasted a total of 1010 ps real time, with the last 1000 ps used in further analysis.

Grid search calculations about ϕ and ψ for each disaccharide fragment of

Gb_4-OEtTMS were generated using the following protocol. First, a series of 24 x 24 geometries for each fragment was created by independent variation of ϕ and ψ in 15 degree steps for each disaccharide fragment. Each structure was then minimised using the quasi-Newton-Raphson algorithm until the maximum derivative was less then 0.4 $kJmol^{-1}Å^{-1}$. The values of ϕ and ψ were constrained to their initial values with a force constant of 210 $kJmol^{-1}rad^{-2}$ during the minimisation. Each gridsearch was then presented as a contour plot using the Biosym CONTOUR utility.

Theoretical time-averaged rotating frame Overhauser effects (ROEs) were computed using an in-house written full-matrix relaxation program using an r^{-6} formalism for simulations at 300K and $(r^{-3})^2$ formalism coupled with exchange data for hydroxyl protons, for simulations at 258K.

Results and Discussion

Proton and Carbon Assignments

Initial assignments were made for Gb_4-OEtTMS in D_2O at 300K (*Table 3.1*) using COSY data (*Figure 3.2*). The αGal(**III**)-H1 proton resonance was easily identified by its characteristic downfield shift and small coupling constant (4.91 ppm J_{H1H2} ~1 Hz). The coupling pattern is very well resolved for this residue with both the **III** H4 (4.25 ppm) and **III** H5 (4.37 ppm) appearing well downfield of other ring protons and the **III** H6 pro-S and pro-R have degenerate shifts (3.69 ppm). The three remaining residues were not so easily assigned by the COSY spectrum alone. The βGlc (**I**) spin system was identified by the characteristic **I** H2 upfield shift

(2.28 ppm) and was completely assigned by means of TOCSY data (data not shown). It's apparent from table 3.1 that the remaining H3/H4/H5/H6/H6′ resonances all lie in a crowded spectral envelope spanning 3.9-3.5 ppm. J-coupling constants for galactose H4 and H5 ($^3J_{H4-H5}$) are smaller than the linewidth (~ 1 Hz), thus, no COSY-type cross-peaks are observed connecting H4 to H5 in residues GalNAc (**IV**), **III**, βGal (**II**). Additional experiments were therefore required to assign the remaining resonances. The TQF-COSY (*Figure 3.3*) selectively detects the H5-H6/H6′ resonances and proved invaluable in assigning these complex spin systems. The remaining **IV** H4 and **II** H4 and ^{13}C resonances were finally assigned using the^{13}C- ^1H HMQC spectrum (*Figure 3.4*).

Assignments made for the Gb_4-OEtTMS in H_2O/acetone-d_6 at 258K were achieved using an ES-TOCSY spectrum (data not shown). Again full assignments of **III** and **I** spin systems could be made directly from the H1 rows of these spins in the TOCSY spectrum. H1-H3 protons of residues **IV** and **II** could be assigned directly from the TOCSY spectrum. However, the H4/H5/H6/H6′ spins could only be assigned based on assignments made in the TOCSY acquired in D_2O. Assignment of hydroxyl protons was achieved by identifying cross-peaks to the anomeric protons. OH6s were assigned via TOCSY cross-peaks to H5 protons. The NH proton was easily assigned due to its characteristic downfield shift (8.47 ppm).

Table 3.1: Proton Assignments for Gb_4-OEtTMS in D_2O at 300K. Referenced to TSP (δ 0 at 300K).

Residue	GalNAc	αGal	βGal	βGlc
Proton	IV	III	II	I
H1	4.63	4.91	4.54	4.49
H2	3.94	3.89	3.58	3.28
H3	3.75	3.96	3.74	3.62
H4	3.93	4.25	4.03	3.66
H5	3.66	4.37	3.79	3.58
H6 pro-R	3.80	3.69	3.92	3.82
H6 pro-S	3.76	3.69	3.84	3.98
CH_3	2.12	-	-	-
C'H	3.77			
C'H'	4.03			
C''H	0.99			
C''H'	1.09			
$Si(CH_3)_3$	0.01			

Table 3.2: Carbon Assignments for Gb_4-OEtTMS in D_2O at 300K. Referenced indirectly to TSP (δ 0).

Residue	GalNAc	αGal	βGal	βGlc
Carbon	IV	III	II	I
C1	104.62	104.75	104.67	102.68
C2	53.85	68.94	76.04	74.32
C3	72.26	80.09	73.75	80.14
C4	69.04	70.31	78.59	75.90
C5	76.22	71.75	76.70	72.33
C6	62.44	61.90	61.79	61.741

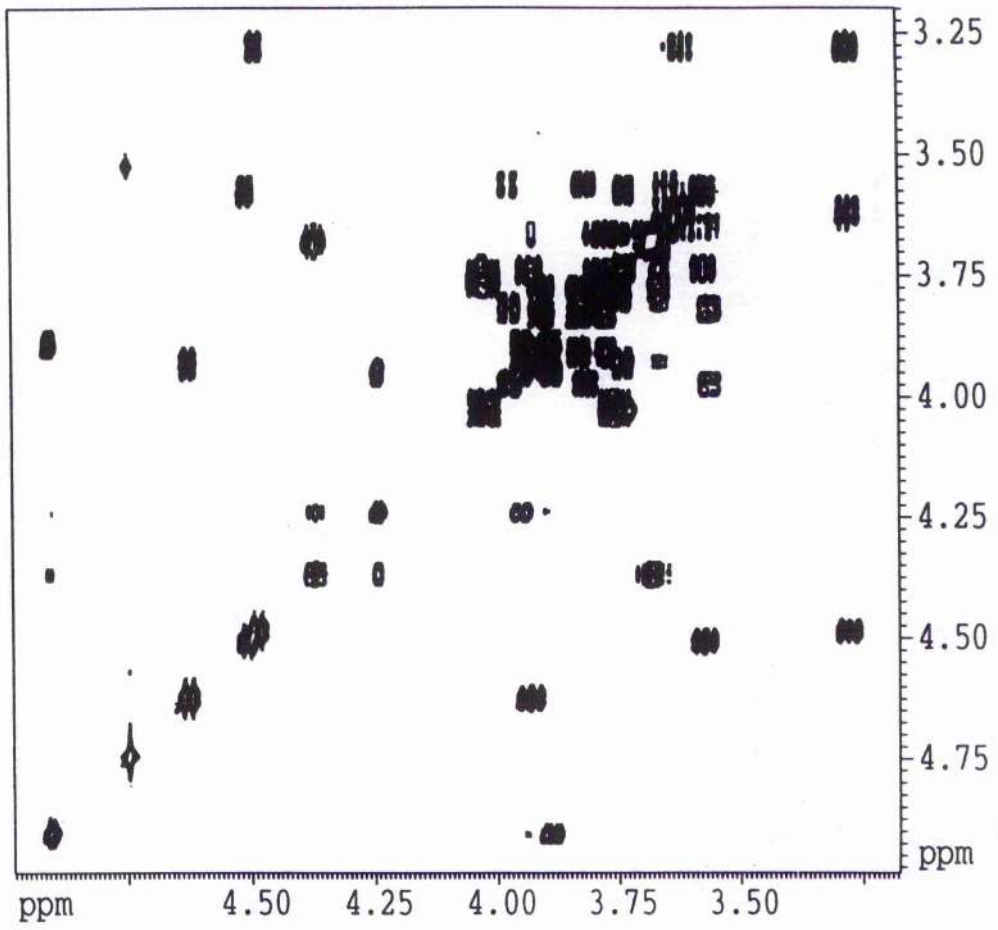


Figure 3.2: 2D COSY Spectrum of Gb_4-OEtTMS in D_2O at 300K.

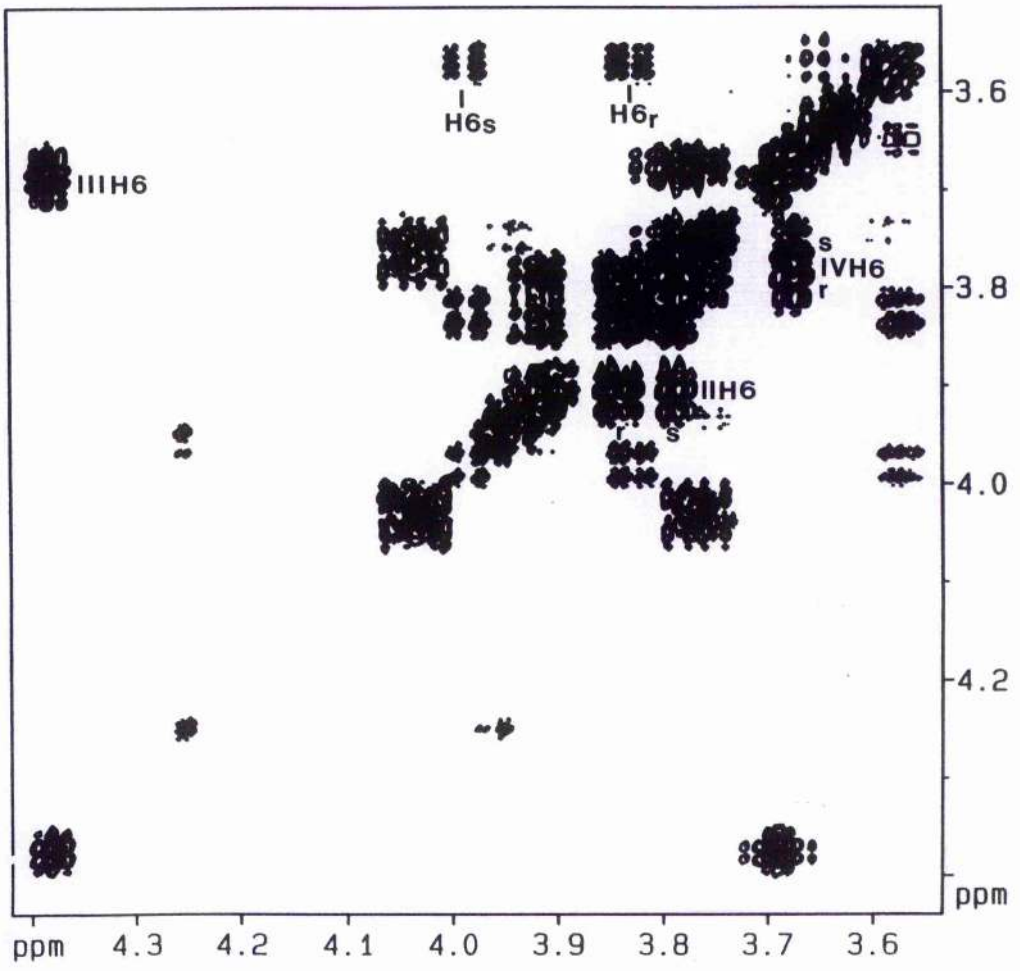


Figure 3.3: 2D TQF-COSY spectrum of Gb_4-OEtTMS in D_2O at 300K.

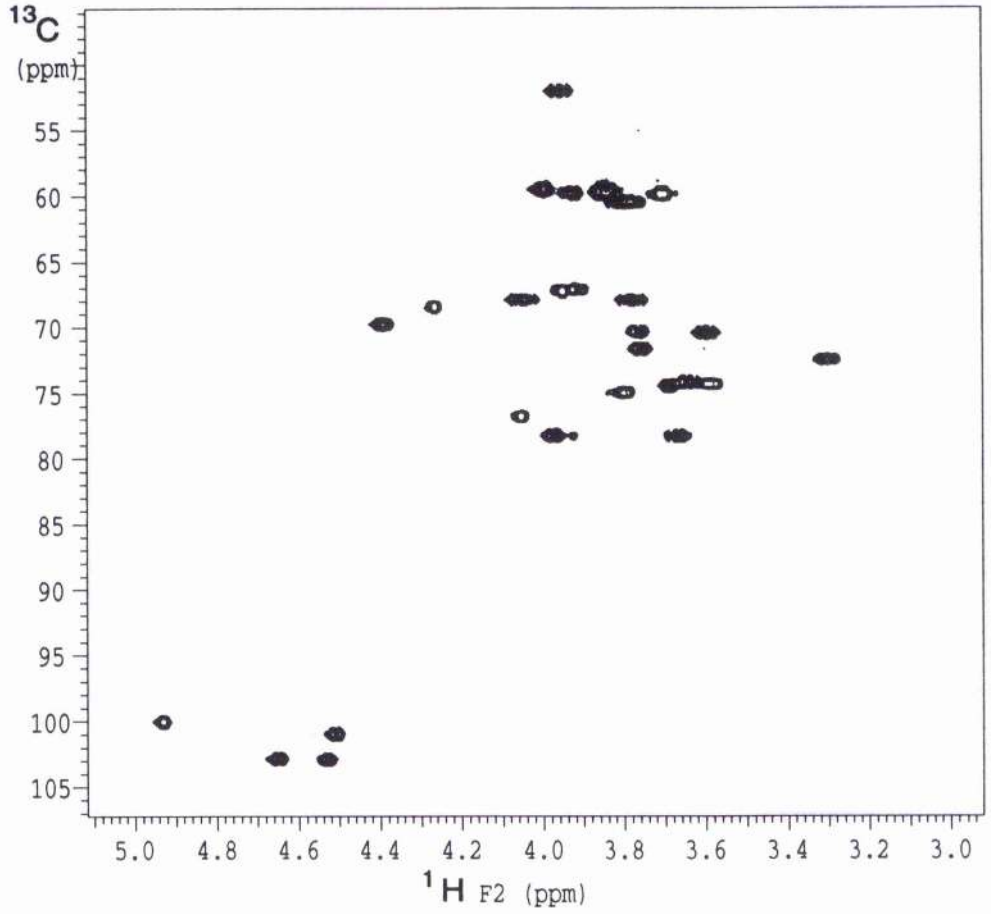


Figure 3.4: Heteronuclear ^{13}C- ^{1}H correlation (HMQC) spectrum of Gb_4-OEtTMS in D_2O at 300K.

Table 3.3: Proton Assignments for Gb_4-OEtTMS at 258 K in 15% Acetone-d_6 Referenced to TSP (δ 0 at 258 K).

Residue	GalNAc	αGal	βGal	βGlc
Proton	IV	III	II	I
H1	4.61	4.90	4.52	4.51
H2	3.96	3.91	3.62	3.28
H3	3.77	3.98	3.77	3.63
H4	3.94	4.29	4.03	3.68
H5	3.68	4.44	3.75	3.58
H6 pro-R	3.86	3.88	3.72	3.86
H6 pro-S	3.81	3.88	3.80	4.00
CH_3	1.98	–	–	–
OH2	–	6.10	6.59	6.62
OH3	6.18	–	5.90	5.98
OH4	6.03	5.94	–	–
OH6	5.82	5.82	6.23	6.27
NH	8.47	–	–	–

ROESY Assignments

Observed inter-residue ROEs between C-linked protons and their ROE ratios (normalised to an internal intra-residue cross-peak volume) are shown in table 3.4. Figure 3.5 shows the reporter regions where all of the observed ROE connectivities are found. Measurement of cross-peak volumes for ROEs to **II** H1 and **I** H1 is complicated due to spectral overlap. Crosspeak volumes for residues **II** and **I** were estimated by measuring one half of the crosspeak multiplet in *F1* and doubling the volume.

NOESY Assignments

Figure 3.6 shows an enlarged region of the ES-NOESY spectrum showing NOE contacts between hydroxyl and non-exchangeable protons. All of these NOE contacts have been observed in ROESY spectra obtained for the globoside dissolved in Me_2SO and Me_2NCHO. However, some of the ROEs observed in these organic solvents are not observed in the ES-

NOESY spectra of Gb_4-OEtTMS. These missing NOE contacts are between exchangeable protons, and dynamic exchange with the bulk solvent may have attenuated these cross-peaks in H_2O.

Table 3.4: Measured ROE values and distance restraints (Å) for initial modelling studies of Gb_4-OEtTMS.

Proton Pair	Relative ROE	r(Å[a])	Restraint
IV H1 - **III** H3	1.24[b]	2.27	m
IV H1 - **III** H4	0.18[b]	3.20	w
IV H1 - **III** H2	*0.81*[b,g]	*2.43*[h] *2.8*[i]	w
III H1 - **II** H4	1.22[c]	2.30	m
III H1 - **II** H6 $_{proR}$	0.15[c]	3.28	w
III H5 - **II** H2	0.61[d]	2.82	w
III H5 - **II** H4	0.11[d]	3.70	w
II H1 - **I** H4	1.09[e,f]	2.55	s
II H1 - **I** H6 $_{proS}$	0.19[e]	3.05	w

[a]Approximate effective ^1H-^1H distance, based upon a simple ROE ratio calculation which neglects the influence of internal motions and anisotropic reorientation.

[b]Reference ROE **IV** H1-H5, reference distance 2.35 Å.

[c]Reference ROE **III** H1-H2, reference distance 2.40 Å.

[d]Reference ROE **II** H5-H2; reference distance 2.35 Å

[e]Reference ROE **I** H1-H5, reference distance 2.38 Å.

[f]Volume estimated due to spectral overlap with other crosspeaks (discussed in the text).

[g]A ROE cross-peak is observed (*Figure* 3.5), however its volume is enlarged due to the presence of strong coupling.

[h]Distance calculated with rf carrier frequency placed downfield of the proton spectrum.

[i]Distance calculated with rf carrier frequency placed upfield of the proton spectrum

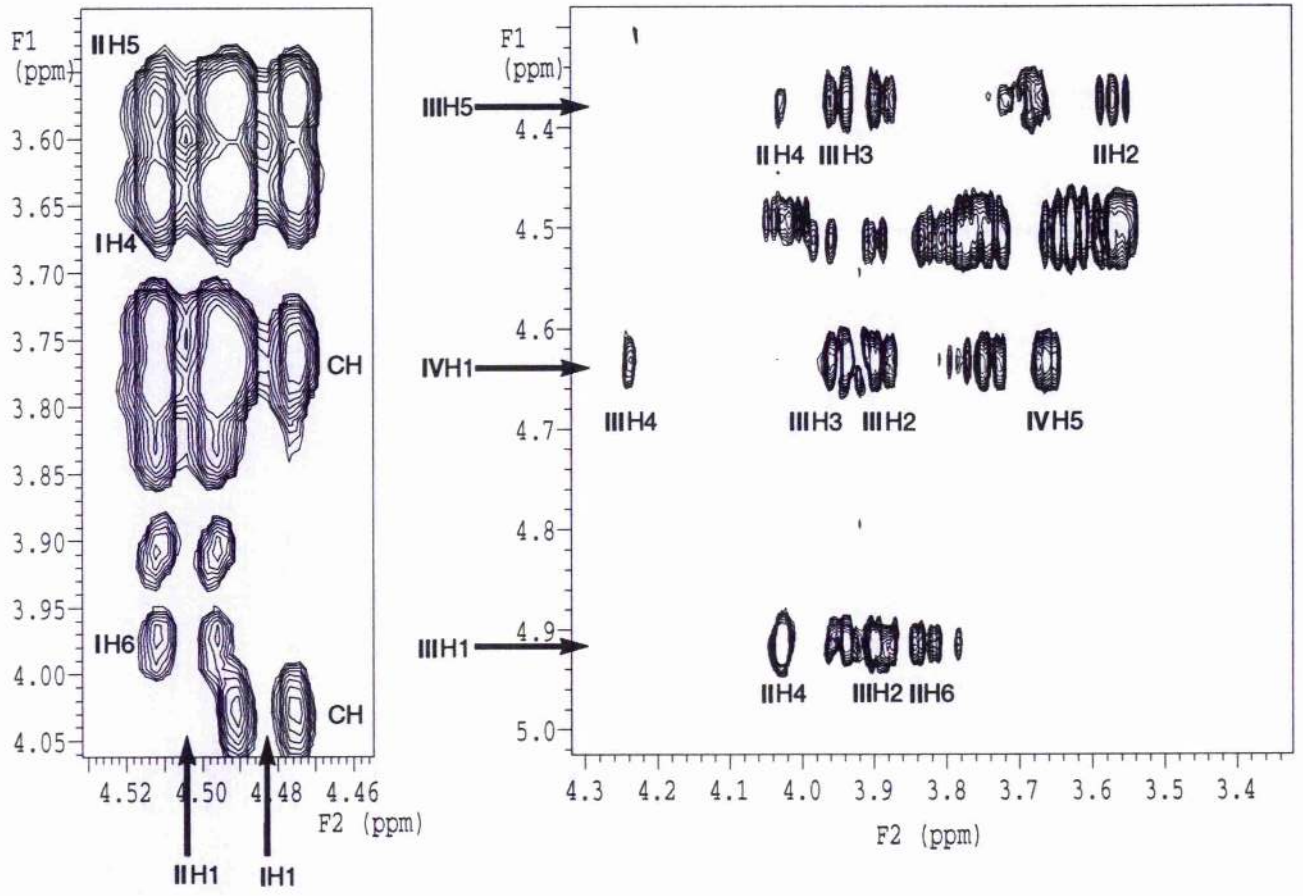


Figure 3.5: Extracts from the 2D ROESY spectrum of Gb_4-OEtTMS in D_2O at 300K with a mixing time of 300 ms. All inter-residue ROEs and intra-residue ROEs used for calibration purposes are highlighted.

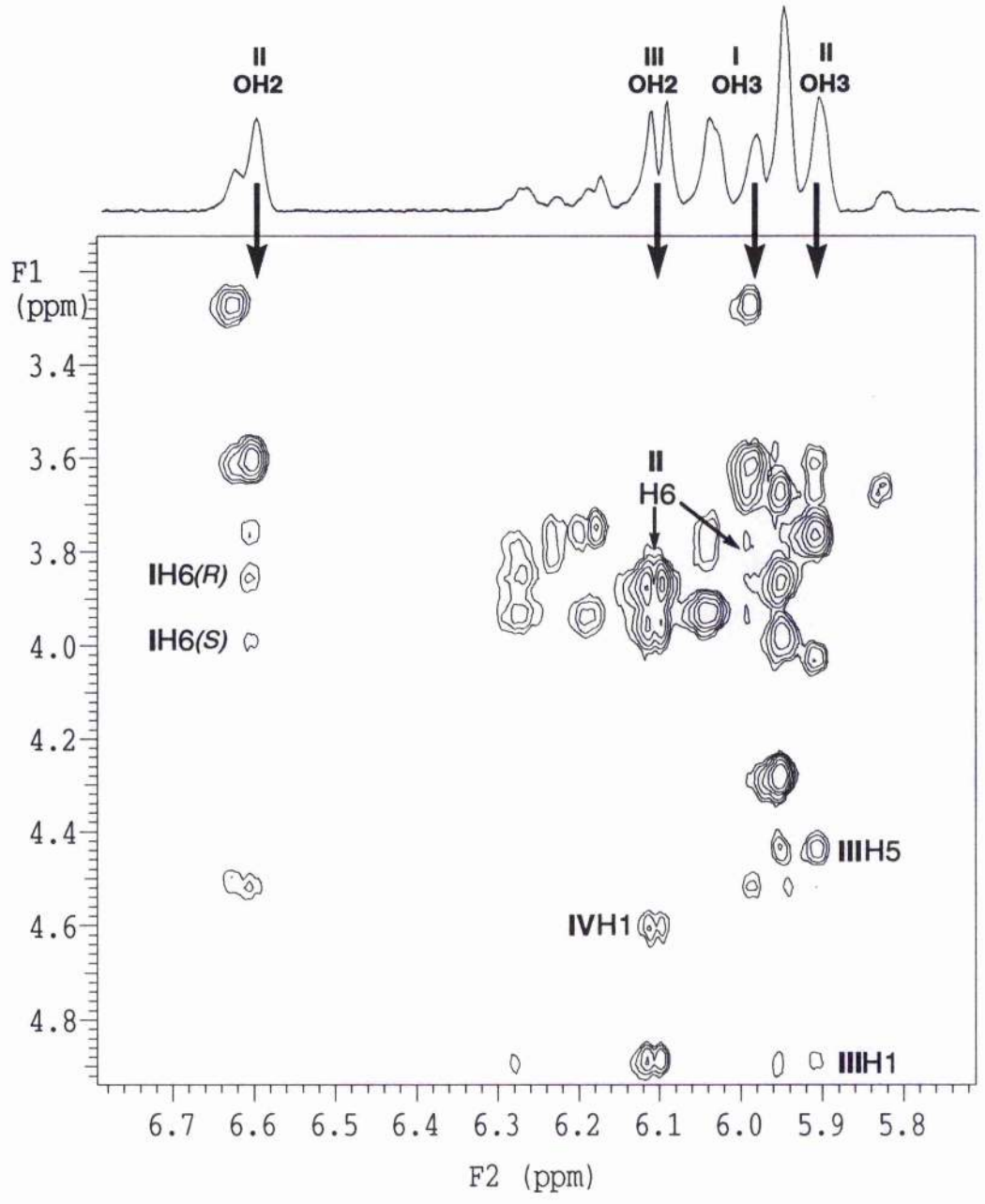


Figure 3.6: Extract from the ES-NOESY spectrum of Gb_4-OEtTMS in H_2O containing 15% acetone-d_6 at 258K. Additional NOE used in low temperature dynamics simulations (except the IV NH - III OH2) are highlighted.

Conformation of the Entire Oligosaccharide Chain

To modelling the 3D structure, the AMBER forcefield with additional parameters derived for carbohydrates (Homans 1990) was used for molecular mechanics and dynamics simulations. Twenty pseudo-random starting geometries followed by restrained simulated annealing and minimisation to give a family(ies) of minimised structures consistent with the NOE distance restraints. Simulations based on distance restraints derived from ROE contacts observed in D_2O annealed into two families (*See table 3.7, Figure 3.5*) about the **IV - III** and the **III - I** linkages, and annealed into multiple families about the **II - I** glycosidic linkage. The ϕ,ψ values for **III - II** and **II - I** have been reported previously for the intact globoside. However, a previously unreported solution conformer about the **IV - III** linkage has been found by this modelling protocol. The conformer (ϕ,ψ~ -20, -40) has not been observed for the intact globoside but has been reported for Galβ1-3GalNAc and Galβ1-3Gal linkages (Weller 1994). The overall conformation of the sugar at the global minimum is an "L" shape with the **IV** *N*-acetyl-group resting in a cleft formed by the **III**, **II**, and **I** groups whilst the local minimum which is approx. 8 kJ mol^{-1} has a similar overall conformation apart from the *N*-acetyl group which is oriented out into solution (*See Figure 3.7B*).

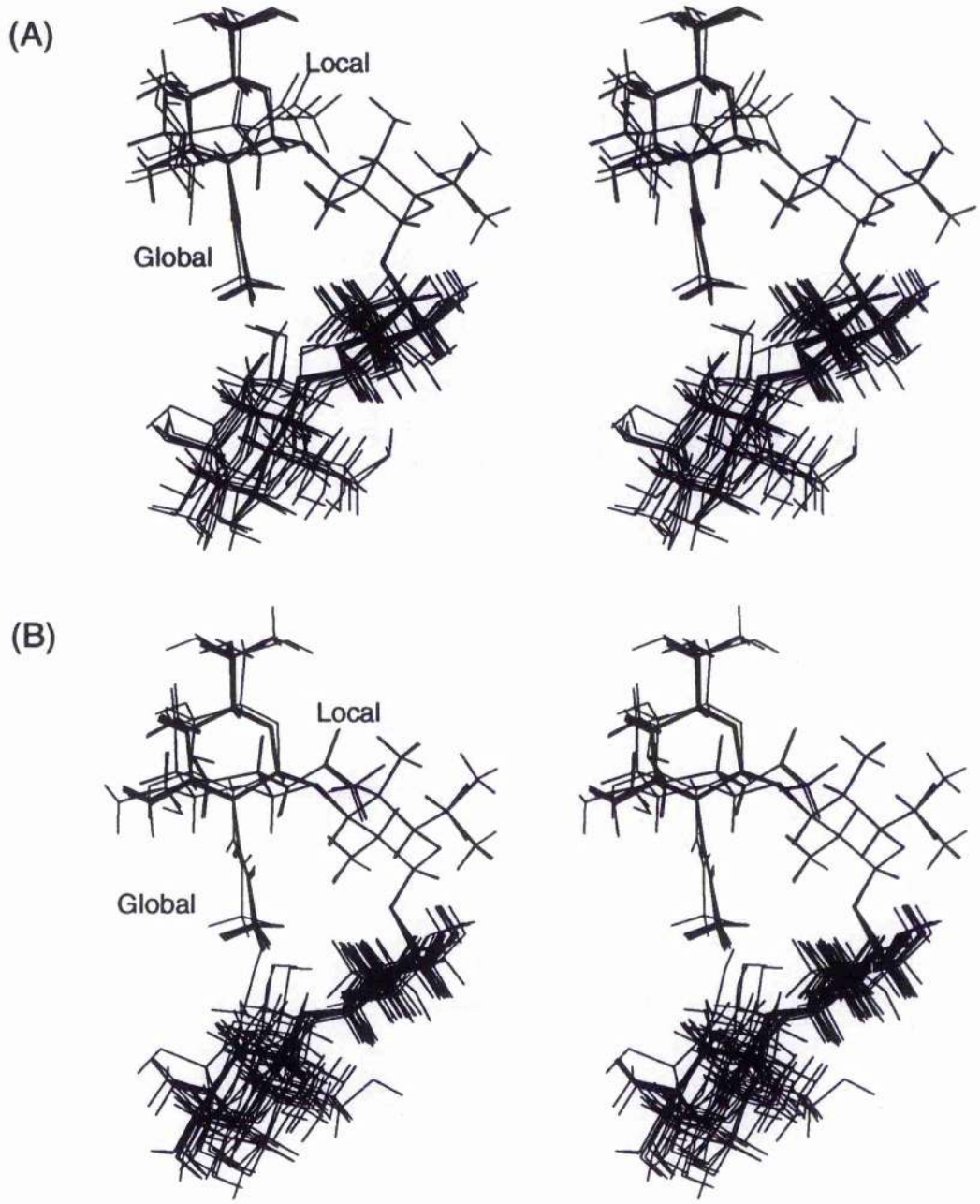


Figure 3.7: Structures from restrained simulated annealing modelling (A): ROESY derived structures. (B): ES-NOESY derived structures.

Table 3.5: Energies and angles of structures from the simulated annealing process which satisfied all of the distance restraints obtained from (A) ROESY data at 300K and (B) ES-NOESY data at 258K. Angles in bold type are representative structures used in restrained dynamics simulations.

(A)	IV - III		III - II		II - I		Energy
	ϕ	ψ	ϕ	ψ	ϕ	ψ	kJ mol^{-1}
Global	**-22.0**	**-46.0**	**-37.0**	**-4.9**	**-21.5**	**-3.3**	**0.0**
2	-21.9	-46.3	-26.0	12.8	30.5	-10.8	4.1
3	-22.2	-45.7	-37.0	- 3.9	-22.8	-27.3	2.1
4	-24.2	-47.0	-32.2	0.7	-21.0	-26.8	2.1
5	-20.5	-45.4	-40.4	9.4	43.9	- 6.4	3.7
6	-24.7	-46.0	-36.0	11.8	34.2	- 4.5	3.8
7	-23.2	-45.2	-37.3	9.6	33.9	- 3.1	3.7
8	-21.9	-44.9	-35.8	0.4	32.5	- 6.9	3.7
9	-24.6	-46.8	-31.7	- 0.1	-25.6	-26.5	3.8
Local	**25.7**	**20.8**	**-23.7**	**14.3**	**51.6**	**- 0.1**	**5.9**
11	20.0	26.1	16.7	25.3	-0.8	-26.8	10.9
12	20.0	25.3	-24.5	15.8	51.0	- 0.7	9.2
13	19.7	24.7	-45.0	- 6.9	-22.2	-24.2	10.9
14	20.0	24.8	-27.9	11.9	45.2	0.4	12.6
(B)							
Global	**-22.3**	**-44.7**	**-43.8**	**1.1**	**-21.5**	**-26.8**	**0.0**
2	-21.5	-44.7	-35.3	0.5	32.8	-3.3	3.8
3	-22.4	-44.3	-40.3	5.3	32.2	-4.4	2.1
4	-19.2	-47.8	-32.6	14.9	44.4	- 6.9	3.4
5	-23.5	-46.4	-32.4	-1.3	-20.9	-26.5	2.9
6	-21.7	-45.2	-34.5	5.4	32.6	- 4.0	3.4
7	-26.7	-47.6	-33.7	14.0	-15.4	-27.5	5.5
8	-23.6	-45.9	-38.4	9.5	31.3	- 7.2	4.6
9	-21.4	-44.8	-34.1	7.4	31.3	- 6.1	4.5
10	-22.8	-45.3	-33.1	3.1	4.4	-25.1	1.8
11	-21.0	-45.7	-27.8	12.1	32.6	- 5.4	4.5
12	-16.4	-45.4	-26.1	14.1	46.0	1.3	5.0
Local	**19.8**	**25.3**	**-26.5**	**14.5**	**-1.3**	**-26.5**	**5.0**
14	26.7	26.2	-28.7	10.7	80.6	- 8.4	8.8

This novel solution minima about **IV** - **III** is proposed on the basis of restrained simulated annealing with only 3 inter-residue ROE contacts describing the conformation of the linkage and it could be argued that this novel conformation is perhaps due to an artifact of the modelling protocol. To test the validity of this new solution conformation more inter-residue data are required. As additional restraints are used in the dynamics simulations the likelihood that the final conformer is due to an artifact of the forcefield is reduced. NOE contacts between hydroxyl and C-linked protons are a potentially rich source of additional restraints. Hydroxyl protons have much greater conformational freedom and extend further out into solution compared to C-linked protons, however, the observation of hydroxyl protons is not a trivial problem. Significant advances in this regard has been the work of Poppe and van Halbeek, who have demonstrated that it is possible to observe the exchangeable (-OH and -NH) proton resonances of oligosaccharides in acetone/H_2O (Poppe and van Halbeek, 1991) or in supercooled H_2O solution (Poppe and van Halbeek, 1994). The rate of exchange at these temperatures (~ 255 K) is sufficiently slow that narrow linewidths are obtained, thus allowing straightforward assignment of these resonances and measurement of NOEs to and from them. In this manner, the number of usable conformational restraints is substantially increased (Poppe *et. al.*, 1992).

Figure 3.6 shows the reporter region of ES-NOESY spectra of Gb_4-OEtTMS dissolved in H_2O/acetone-d_6 at low temperatures (-15°C) and highlights inter-residue NOE contacts. On average there are 3 additional restraints per glycosidic linkage. The solution conformations calculated from ROE restraints measured in D_2O were also consistent with NOE contacts observed between C-linked protons and exchangeable protons (*See*

Figure 3.5A). Simulated dynamical annealing of Gb_4 with NOE restraints derived from NOESY data obtained in H_2O/acetone-d_6 resulted in the same two families of solution conformers observed for simulations using non-exchangeable ROE restraints only (*See Figure 3.5B*).

Solution Dynamics of G_{b4}

It has long been established that in mobile species such as oligosaccharides proton-proton distances derived from NOE or ROE data are unlikely to represent the single or true solution conformer but a r^{-6} weighted average of all of the conformers accessed in solution (often referred to as a 'virtual' conformation (Carver *et al.* 1989)). To simulate in part the dynamical behaviour of the glycan in solution, restrained molecular dynamics simulations were performed. Spectral data, for example, ROEs can be back-calculated by means of a full relaxation matrix algorithm to check their validity. Two conformers were chosen from the annealed structures with both the lowest energy 'global' minimum conformation and a 'local' energy conformation (defined as the lowest energy structure occupying the alternate energy minima about the **IV - III** and **III - II** linkages) chosen as starting structures for one ns of restrained simulated annealing.

Solution Dynamics of the GalNAcβ1-3Galα (IV - III) Linkage

Effective proton-proton distances for **IV** H1 - **III** H3/H4 (listed in table 3.4) are similar to those that have been reported previously by Poppe and co-workers (1990) for the intact globoside. However, in this study the **IV** H1 - **III** H2 ROE cross-peak intensity is complexed by the presence of relayed transfer effects (Keeler, *etal* 1987) operative on **III** H2 and **III** H3 (J ~ 11 Hz

$\Delta\upsilon$ ~ 37 Hz), the latter being close in space to the **IV** H1 proton. This results in an unrealistically short proton-proton distance between **IV** H1 - **III** H2 (2.43Å). This complex cross-peak pattern is repeated in every row of the ROESY spectrum (*Figure 3.5*) and may therefore be a 'false' ROE enhancement. So-called 'false' ROE enhancements may be detected by shifting the frequency offset and looking for changes in cross-peak intensity (Neuhaus 1990). This methodology was applied to Gb_4-OEtTMS by performing two different ROESY spectra identical in every way except that the transmitter offset was placed at opposite ends of the spectrum. Recalculating the effective proton-proton distances for all the inter-residue ROEs resulted in almost identical ($\Delta r = \pm 0.01$Å) back calculated distances for all pairs of spins except for **IV**H1 - **III**H2 which shifted significantly ($\Delta r = 0.4$Å). Due to this ambiguity, the distance restraint set for **IV** H1 - **III** H2 was set to weak.

Simulated annealing resulted in two families of minimum energy conformations about this linkage (*Figure 3.7A* and Table 3.5A). One family annealed to a conformation similar to that reported for the intact globoside (ϕ,ψ~ 25°, 25°)(Scarsdale *et al.* 1986) the other major family (which included the lowest energy conformer) sits in a novel energy minima consistent with the ROE distance restraints (ϕ,ψ~ -20°, -40°). Back calculated ROE intensities for this linkage require careful analysis (See Table 3.6).

Table 3.6: Measured ROE values and back-calculated NMR data from 1 ns dynamics simulations.

	Experimental	Simulated Relative ROEs			
Proton Pair	ROE	Restraint	Unrestrained[a]	Global Min.	Local Min.
IV H1 - III H3	1.24	m	1.13	0.77	0.98
IV H1 - III H4	0.18	w	0.36	0.36	0.47
III H1 - II H4	1.22	m	1.24	1.27	1.26
III H1 - II H6 proR	0.15	w	0.22	0.16	0.19
III H5 - II H2	0.61	w	0.92	0.74	0.64
III H5 - II H4	0.11	w	0.07	0.06	0.06
II H1 - I H4	1.09	s	0.55	1.15	1.21
II H1 - I H6 proS	0.19	w	0.03	0.21	0.16

[a]Unrestrained dynamics simulation performed on the global energy minumum conformation.

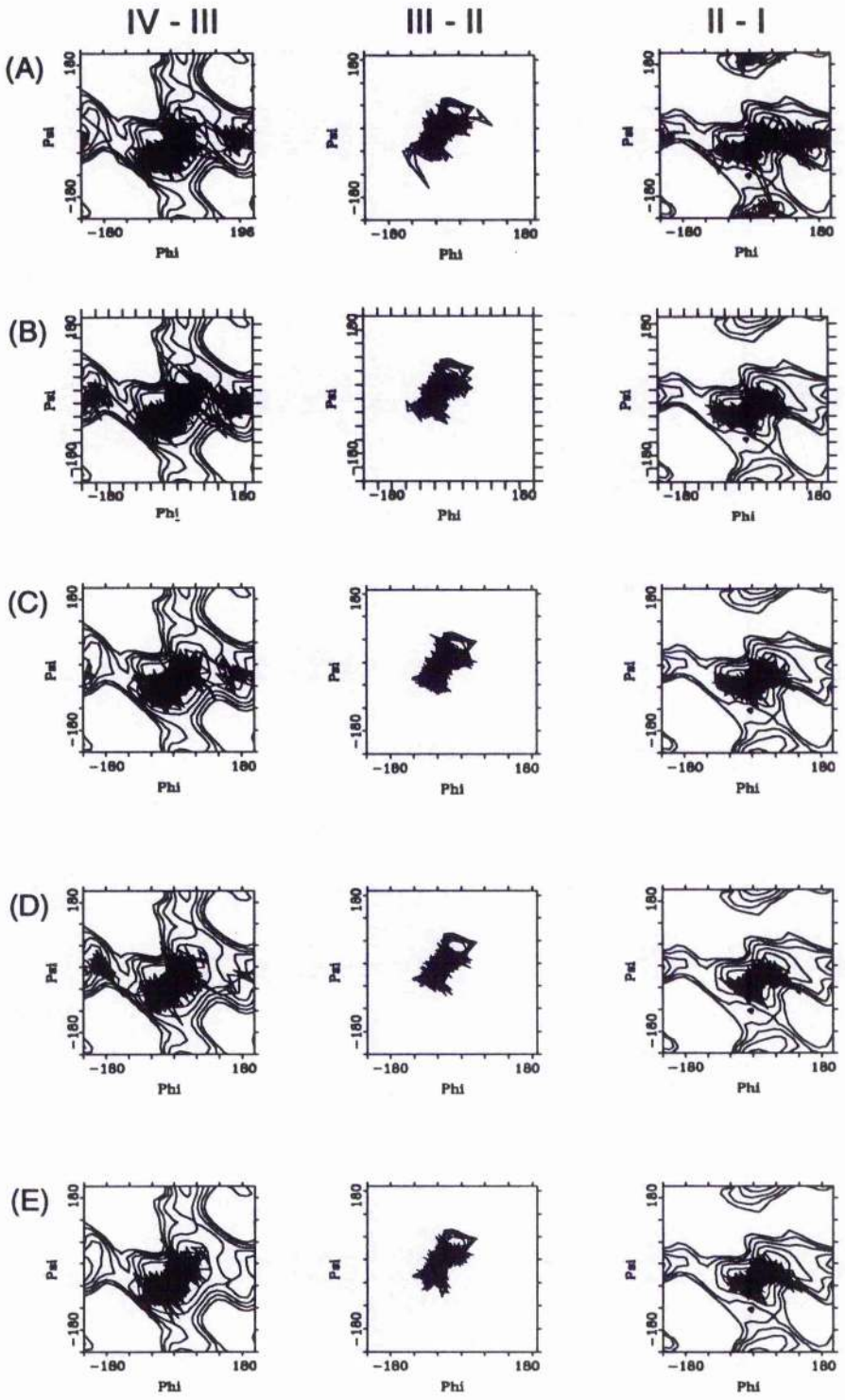


Figure 3.8: Phi vs. Psi trajectory plots for Gb_4-OEtTMS. (A): Unrestrained dynamics at 300K. Restrained molecular dynamics at 300K with ROE restraints starting from (B): global minima (C): local minima. Restrained molecular dynamics at 258K with NOE restraints starting from (D): global minima (E): local minima.

Theoretical analysis of unrestrained dynamics simulations (simulations which show the maximum extent of conformational mobility consistent with the force-field) give back calculated relative ROEs which are within experimental error (within a factor of two, as discussed in detail in Rutherford, T. J. and Homans, S. W. (1994)). The restrained dynamics simulations (simulations which show the minimum conformational mobility consistent with NMR data) on the global energy minimum also result in a reasonable match with experimental relative ROEs. Analysis of the trajectories for these two simulations indicate that there are rapid transitions between the local and global energy minima. However, there are also frequent transitions between the broad central potential energy well and an additional energy minima located at ϕ,ψ -175°, 0°. It is interesting to note that the dynamics simulation performed on the local energy minimum, which over estimated relative ROE for the **IV** H1 - **III** H4, had rather fewer transitions into the ϕ,ψ -175°, 0° minima. This would suggest that the **IV** - **III** linkage is more accurately described by a more flexible model for solution behaviour. The time-averaged conformation for this linkage (ϕ,ψ ~30°, -15°) lies closer to the local minimum than the global minimum. Recently, a model of the bound conformation of Gb_4 in complex with VT2e has been proposed (Nyholm *et al.* 1996b). These modelling studies predict that the **IV** - **III** linkage will adopt a (ϕ,ψ ~35°, -29°) conformation when bound to VT2e, a conformation which is similar to the time-averaged conformation predicted in this present study.

Solution Dynamics of the Galα1-4Galβ (III - II) Linkage

In common with the unrestrained simulation of the **IV - III** linkage back calculated ROE parameters from the unrestrained dynamics trajectories are in excellent agreement with experimental data. Dynamics simulations performed with ROE restraints do result in a slight improvement of the theoretical relative ROEs. Analysis of the the potential energy surface and trajectory plots (*See Figure 3.8A*) indicate that this linkage is sterically crowded and the molecule can only access a small region of conformational space. It is clear that the alpha linkage is responsible for holding the molecule in its time-average "L" shape.

Solution Dynamics of the Galβ1-4Glcβ (II - I) Linkage

This linkage occurs in almost every series of glycosphingolipids. The Galβ1-4Glc (**II - I**) linkage is the only linkage that is significantly improved by the inclusion of atomic constraints. Unrestrained dynamics simulations overestimate the degree of torsional freedom predicted for this linkage. Restrained dynamics simulations are in excellent agreement with experimental data. As with previous modelling studies, pronounced conformational freedom is predicted for this linkage.

Solution Dynamics and Structure of Gb_4-OEtTMS in H_2O/acetone-d_6 Solutions

The extraction of quantitative distance information from crosspeak intensities involving exchangeable protons is not straightforward. First, as discussed by James and co-workers (Liu *et. al.*, 1993), a loss of crosspeak intensity arises from the exchange of magnetisation with solvent water. This is effectively a leakage process which affects only the diagonal

elements of the relaxation matrix (ρ_I^*), and can be accounted for in a full relaxation matrix calculation of the ROESY or NOESY intensities provided the exchange rates are known. The latter can be measured in a straightforward manner using, for example, the approach described by Adams and Lerner (1992). A second problem concerns the extent of saturation of the water magnetisation. If the water magnetisation is saturated prior to acquisition, substantial saturation transfer can occur from the bulk solvent (which is in vast excess) to the solute during the acquisition period and relaxation delay, and can result in a substantial reduction in resonance intensity of exchangeable protons (Li and Montelione, 1993; Grzesiek and Bax, 1993; Stonehouse *et. al.*, 1994). This effect can in principle be eliminated by ensuring that the water magnetisation returns to equilibrium between each transient (Stonehouse *et. al.*, 1994), but this is inefficient since the T_1 of water is often long. An alternative is to employ selective 'flip-back' pulses (Grzesiek and Bax, 1993) or judicious choice of phase cycling (Mori *et. al.*, 1995) together with tailored excitation techniques such as 'WATERGATE' (Sklenár *et. al.*, 1993). In this current study 'Excitation sculpting' (ES) originally developed by Hwang and co-workers (1995) is applied. In this pulse scheme, denoted by A=G_1-S-G_1-G_2-S-G_2, where S is a soft-180$_x$ (water-selective) followed by a hard-180$_x$ pulse, and G_1 and G_2 are gradients satisfying the condition that the application of G_2 does not refocus magnetisation dephased by G_1, transverse magnetisation is returned to its original position by the action of A. The power of this pulse sequence A, is that it can be appended onto almost any homonuclear pulse sequence and water suppression levels of > 30 000-fold have been reported.

Table 3.7: Back-calculated NOEs for additional NOEs observed for Gb_4-OEtTMS in 15% acetone-d_6/ H_2O.

Proton Pair	Relative NOE	Theoretical NOEs Global	Local
IV H1 - **III** OH2	0.15[a]	0.09	0.11
IV NH - **III** OH2	0.09[a,f]	0.08	0.10
III H1 - **II** OH3	0.10[b]	0.10	0.10
III H5 - **II** OH3	0.36[b]	0.32	0.28
II H6(R) - **III** OH2	0.13[b]	0.16	0.10
II H6(R) - **I** OH3	0.13[c]	0.13	0.20
I H6(R) - **II** OH2	0.14[c]	0.32	0.31
I H6(S) - **II** OH2	0.23[c]	0.52	0.51

[a]Reference NOE **IV** H1- H5

[b]Reference NOE **III** H1 - H2

[c]Reference NOE **I** H1 - H2 (estimated volume).

[d]Confirmed as a real NOE from ES-ROESY (data not shown).

Application of NOE restraints between hydroxyl and C-linked protons has not reduced the amount of torsional freedom predicted by the restrained dynamics simulations, although the number of oscillations between the major energy well of the **IV - III** linkage and the ϕ,ψ -175°, 0° minimum is reduced when NOE restraints are included (*Shown in Figure 3.8B*). Theoretical relative NOEs for these additional restraints (which were calculated from full matrix relaxation calculations including exchange terms in model exchange with bulk solvent, see Table 3.8) match well with experimental values (Table 3.7) except for the **II - I** restraints between hydroxymethyl protons and exchangeable protons This inconsistency is perhaps due to insufficient modelling of the rotation about the hydroxymethyl group, also severe spectral overlap in the C-linked region

of the ES-NOESY spectrum made measuring accurate intra-residue NOE crosspeaks for calibration difficult.

Inter-residue H-bonds have been reported for sugars in aqueous solutions (Poppe *et al.* 1992), we therefore investigated spectral data for evidence of possible H-bonds in Gb_4-OEtTMS. H-bonds are usually identified in the following ways: slower exchange rates with the bulk solvent and therefore smaller temperature coefficients, and of course suitable geometric properties. Table 3.8 lists the exchange rates for all but the amide proton (which exchanges with the bulk solvent on a much slower time scale). It is clear that the **III**OH2 proton exchanges with the bulk solvent at a relatively slower rate suggesting that this group may be involved in H-bonding, possibly, with the **IV** carbonyl oxygen as an NOE contact is observed between **IV**NH group and **III**OH2. Several hydrogen bonds between the **IV** - **III** groups have proposed for the for intact globoside in organic solvents: these include **IV**NH - **III** OH2, and **IV** CO - **III** OH4/OH2. J-coupling constant for the NH group were measured to be 10.3 Hz in H_2O. The $^3J_{H2NH}$ was measured at ~7.6 Hz for the intact globoside in organic solvents, using a suitable Karplus relationship (Karplus, M. 1959) for amide protons (Bystrov *et al.* 1973) it was postulated that the amide group adopts the **IV**H2-C2-N-NH -120° & 140° rotomers in equal populations. The $^3J_{HH}$ value measured in aqueous solutions predicts this group to occupy the -160° conformer only. This change in conformation suggests that there may be an H-bond broken or formed when the molecule is moved from one solvent system to the other. The temperature coefficients (κ) were finally measured by acquiring several ES-^{1}H spectra from -17° to -8°C (listed in Table 3.8). It is clear that the **III**

OH2 group does have the lowest κ value for hydroxyls, however, it is of the same order of magnitude as all the other OH groups so we therefore have to conclude that there is little evidence of long-term H-bond formation in Gb_4-OEtTMS in aqueous solution.

Table 3.8: Exchange rates at 258K and temperature coefficients for exchangeable protons.

Residue	Nucleus	$k_{exchange}$[a]	κ[b]
IV	NH	-	-9
	OH3	23.6	-14
	OH4	18.6	-13
	OH6	35.5	-17
III	OH2	8.4	-10
	OH4	11.8	-17
	OH6	35.2	-17
II	OH2	15.9	-11
	OH3	12.3	-12
	OH6	35.2	-14
I	OH2	16.5	-15
	OH3	12.7	-10
	OH6	21.3	-13

[a] $k_{exchange}$ values measured in s^{-1} measured at 258K.
[b] Temperature coefficients, κ, (-10^3 ppm°C^{-1}).

3.4 Conclusions

Conventional ROE contacts between C-linked protons of Gb_4-OEtTMS were complemented by more numerous contacts involving amido and hydroxyl protons which were used in restrained dynamical simulated annealing, and molecular dynamics simulations. A novel global

minimum about the GalNAcβ1-3Galα (**IV- III**) linkage was proposed for, and dynamics simulations demonstrate significant flexibility for the molecule. The lowest energy solution structure is different from the proposed structure of Gb_4 in complex with pig edema toxin. In common with other carbohydrate-protein interactions (Bourne *et. al.*, 1990a,b; Cygler *et. al*, 1991; Imberty *et. al.*, 1991 and Low, 1996), VT2e does not appear to bind the lowest energy structure of Gb_4. It is remarkable, however, that the theoretical bound conformation of Gb_4 in complex with VT2e (Nyholm, *et al.* 1996a) is very similar to the time-averaged solution conformer predicted by restrained MD simulations. The low affinity VT-E has for Gb_4-OEtTMS may be used to an advantage as it would allow the experimental study of the bound state conformation of Gb_4-OEtTMS in complex with pig edema toxin through the observation of transferred NOEs (TRNOEs). This is of vital interest as it can be used to test the validity of the theoretical model.

Chapter 4

Synthesis of Carbon-13 Enriched Oligosaccharides

Abstract

The application of stable isotope enriched (^{13}C, ^{15}N) amino acids coupled with the development of multidimensional heteronuclear experiments has greatly facilitated the study of protein structure and dynamics by NMR spectroscopy. Their use has dramatically aided the assignment process of large proteins (up to 30 kDa). Application of heteronuclear NMR techniques to carbohydrates has so far been limited by the availability of uniformly carbon-13 enriched monosaccharides, and efficient methods of coupling these sugars. Here we report alternative chemical and chemoenzymatic syntheses of *N*-acetyl [U-^{13}C] mannosamine (ManNAc), a substrate required for the enzymatic synthesis of *N*-acetyl neuraminic acid (Neu5Ac), a moiety involved in a variety of important biological recognition events. The coupling of [U-^{13}C] Neu5Ac to enzymatically synthesised *N*-acetyl [U-^{13}C] lactosamine (Galβ1-4GlcNAc) to produce α2,3 sialyl *N*-acetyl lactosamine, a precursor required for the synthesis of sialyl Lewis^x antigen is described. [U-^{13}C] Lactose (Galβ1-4Glc) was also [U-^{13}C] sialylated to form the oligosaccharide moiety of the core ganglioside G_{M3} (Neu5Acα2-3Galβ1-4Glc).

4.1 Introduction

NMR Investigations of ^{13}C Enriched Oligosaccharides

Historically, the application of carbon-13 enriched oligosaccharides in NMR studies has been limited by the fact that the only uniformly enriched monosaccharide generally available was glucose. Other monosaccharides were usually only available with a single carbon position enriched. In NMR studies of biological systems, access to singly labelled sugars has tended to limit the application of isotope enriched carbohydrates to the determination of thermodynamic parameters. Binding constants (K) and residence times ($\tau_{M'}$) of a ligand in complex with a protein may be calculated by measuring the change in ligand line widths ($\Delta\upsilon$) versus temperature. Repeating these experiments at different ligand concentrations allows the calculation of ΔG°, ΔS°, and ΔH° of complex formation, (Swift and Connick, 1962 and reviewed in Amin and Chatterjee, 1995). Isotope editing of NMR spectra is essential if accurate line widths are to be measured in the presence of overlapping ligand and protein resonances. [1-^{13}C] galactose-labelled *N*-acetyllactosamine (LacNAc) and [*N*-^{13}C-Ac] *N*-acetylgalactosamine (GalNAc) have been used to determine the thermodynamics of complex formation with *Erythrina cristagalli* agglutinin (Berman *et al.* 1985) and *Artocarpus integrifolia* agglutinin (Sastry *et al.* 1988) respectively.

Partially enriched carbohydrate moieties have been used to determine conformation and dynamics. [1-^{13}C] galactose labelled sialyl Lewis^{x} was used to determine the torsional angle ψ of the Gal-GlcNAc linkage

(Ichikawa 1992). The conformation of glycolipids containing [*N*-^{13}C-Ac] GlcNAc (Hare *et al*. 1994), [U-^{13}C] monogalactosylglycerol (Howard and Prestegard 1995), *N*-acetyl [1, 2, 3-^{13}C]-neuraminic acid, *N*-acetyl [1, 2, 3-^{13}C]-neuraminic acid α2,3-lactosylceramide (G_{M3}) (Aubin *et al.* 1993), and [U-^{13}C]-gluco and galactopyranosides (Sanders and Prestegard 1991; Sanders and Prestegard, 1992) were probed in magnetically orientated membranes by measuring dipolar interactions between ^{13}C-^{13}C and ^{1}H-^{13}C spin-pairs and ^{13}C chemical shift anisotropies. The conformation of membrane bound *N*-acetyl [1, 2, 3-^{13}C] neuraminic acid α2,3-lactosylceramide and [*N*-^{13}C-Ac] GlcNAc in complex with wheat-germ agglutinin (WGA) have also been elucidated (Aubin, and Prestegard, 1993; Hare, *et al.* 1994).

The extension of multidimensional heteronuclear NMR experiments to carbohydrates relies on access to [U-^{13}C] oligosaccharides (or oligosaccharides containing some fully enriched monosaccharide residues) in milligram quantities. Yu and co-workers (Yu *et al*. 1993) first demonstrated that two and three dimensional experiments such as HCCH-COSY and HCCH-TOCSY could be tailored to assign the proton and carbon resonances of [U-^{13}C] β-glucans extracted from yeast grown on [U-^{13}C] glucose. Recently, work performed by the Homans' group studying [U-^{13}C] glycoconjugates (Gilhespy-Muskett *et al*. 1994; Harris *et al.* 1996; Low, 1996; Weller, *et al.* 1996) has demonstrated the application of a variety of heteronuclear NMR techniques, that were originally developed for the study of proteins, to the study of carbohydrate

conformation both in free-solution and interacting with proteins. The solution conformation of a glycosteroid estrone-3-[U-^{13}C]-glucuronide (E3G) was determined using NOE data obtained from 2D NOESY-HSQC experiments and comparing theoretically derived coupling constants with experimentally measured $^{3}J_{CH}$ and $^{3}J_{CC}$ values across the glycosidic linkage (Low, 1996). The conformation of E3G in complex with an antibody fragment (Fv E3G myc) raised against it was also elucidated by comparing theoretical ring current shifts with those measured from a HSQC spectrum of the complex. [U-^{13}C] galactose has been enzymatically coupled to the non-reducing end of the biantennary *N*-linked glycan of the glycoprotein FcREA (Gilhespy-Muskett, *et al.* 1994). HSQC experiments identified that the [U-^{13}C] galactose moieties experienced two quite distinct chemical environments in the sample. One galactose residue appeared to be closely associated with the protein, whilst line shapes arising from the other arm unit on the α1-3 arm were very much sharper, suggesting that this galactose was extending out into solution. These data confirmed observations made with the X-ray crystal structure (Deisenhofer, 1976)

A novel approach to obtaining [U-^{13}C] oligosaccharides has been described in (Weller *et al.* 1996). Recombinant human chorionic gonadotropin (hCG) was expressed in mammalian cells (Chinese hamster ovary cells) using fully ^{15}N and ^{13}C enriched media. Mammalian cells, unlike bacterial systems, can express glycosylated glycoproteins. This allowed hCG glycans to be uniformly ^{15}N and ^{13}C enriched and the solution conformation of the *N*-linked carbohydrate of intact [U-^{13}C, ^{15}N] hCG

could then be studied. HCCH-TOCSY experiments were 'tuned' to selectively detect glycan resonances by careful setting of the ^{13}C carrier frequency (between the anomeric and endocyclic carbon shifts) and the application of a modest DIPSI spin-lock field. HNCA experiments were also used to determine the glycan linkage point to the protein. However, the above protocol is not applicable to the synthesis of [U-^{13}C] mono and oligosaccharides on the milligram scale due to the minute quantities and heterogeneity of sugar expressed per litre of media. In order to extend the number of biological systems that may be studied by modern heteronuclear NMR experiments, access to a wider range of [U-^{13}C] monosaccharides is essential.

Published Synthesis of N-Acetyl Neuraminic Acid

As discussed in Chapter 1 *N*-acetyl neuraminic acid is arguably the most important component of many glycoconjugates. Structural investigations of sialyloligosaccharides of glycoproteins and glycolipids have historically been limited by the scarcity of material either from natural sources (Schauer, 1982; Koketsu. *et al.* 1992) or from chemical synthesis. Therefore, an efficient method for the synthesis of large quantities of sialic acid (Yamamoto *et al.* ; Baumberger and Vasella, 1986; Bednarski *et al.* 1987; Simon. *et al.* 1988; Kragl, 1990; Kragl *et al.* 1991; Kragl, 1992; Gordon and Whitesides, 1993; Kragl, *et al.* 1995) and its derivatives (Augé, *et al.* 1990; Kittelmann, *et al.* 1992; Kragl *et al.* 1993; Sparks *et al.* 1993; von Itzstein *et al.* 1993; Fitz *et al.* 1995; Kong and von Itzstein, 1995) and subsequent glycosylation has been the goal of many groups over the last three decades (Ichikawa *et al.* 1991; Ichikawa 1992; Ichikawa *et al.* 1992; Ito and Paulson 1993; Scudder *et al.* 1993). None of above methods has so far

has been directly applicable for the production of Neu5Ac in fully carbon-13 enriched form, due to difficulties in obtaining suitably labelled precursors and chemical synthesis often suffers from poor regio- or stereo-selectivity. The synthesis of uniformly carbon-13 enriched *N*-acetyl lactosamine and lactose moieties has also been limited by access to suitably labelled nucleotide donors and acceptors. The aim of this chapter is to describe the chemoenzymatic synthesis of uniformly labelled substrates required for the synthesis of carbon-13 enriched Neu5Ac, LacNAc and lactose. The enzymatic coupling of labelled Neu5Ac to LacNAc and lactose is also described.

4.2 Materials & Methods

Conventions

In order to simplify the text the following nomenclature will be used when describing labelled compounds: the prefix **[U-^{13}C]** will be used to describe mono or oligosaccharides which are carbon-13 enriched (> 99%) at every carbon position *except* the carbon atoms of an *N*-acetyl moiety, unless otherwise indicated.

All chemicals were purchased from Sigma unless otherwise indicated. All enzymatic reactions were incubated at 37 °C unless otherwise stated.

Purification of Neutral Sugars

Monosaccharides: Bio-gel P4 resin (Bio-Rad) (P4 resin 400 mesh, 117 x 1.5 cm, eluant: H_2O @ 12 mL h^{-1}). Disaccharides: Bio-gel P2 resin (Bio-Rad) (P2 resin fine, 117 x 1.5 cm, eluant: H_2O @ 10 mL h^{-1}).

Purification of Charged Sugars

P2 resin (P2 resin fine, 117 x 1.5 cm, eluant: 0.1 M NH_4HCO_3 @ 8 mL h^{-1}) The detector was an ERMA 7512 refractive index detector.

Synthesis of N-Acetyl [U-^{13}C] Glucosamine (Scheme 1)

The following reactions were carried out in a 10 mm NMR tube, monitored by ^{31}P NMR. [U-^{13}C] Glucose (100 mg) (MSD Isotopes), ATP (250 mg), and L-glutamine (70 mg) were dissolved in 5 mL of 50 mM Tris-HCl, pH 7.4 containing 200 mM $MgCl_2$. The pH of the reaction mixture was readjusted after the addition of ATP with 1M NaOH. Hexokinase (E.C. 2.7.1.1) (10 U), phosphoglucose isomerase (E.C. 5.3.1.9), (50 U) and glucosamine 6-phosphate synthase (4 U)(E.C. 2.1.16) (Badet, B. *et al.* 1987; Dutka-Malen, S. *et al.* 1988) was added. The reaction was left to incubate for 18 hours. [U-^{13}C] glucose and phosphoglucose isomerase have also been substituted by [U-^{13}C] fructose (CK products) (100 mg) in the above reaction scheme.

The reaction mixture was adjusted to pH 9.6 with 1 M NaOH. Alkaline phosphatase (E.C. 3.1.3.1) (100U) was added, the reaction was incubated for 12 hours. The reaction vessel was placed in a boiling water-bath for 5 minutes to destroy all remaining enzyme activity. The final solution was stored at -20 °C until required. Synthesis of [U-^{13}C] glucosamine was confirmed by ^{1}H NMR of the crude reaction mixture.

The [U-^{13}C] glucosamine reaction mixture was saturated with $NaHCO_3$ (s). The reaction vessel was placed in an ice bath and 250 µL of acetic

anhydride was added. Additional aliquots of acetic anhydride were added in 15 minute intervals until all the remaining $NaHCO_3$ had reacted. A partial purification was achieved by passing the reaction mixture through a mixed-bed column (Dowex AG50 X12 over Dowex AG3 X4) and washing with water as an eluant. Additional purification of [U-^{13}C] *N*-acetyl glucosamine was achieved by Bio-gel P4 chromatography. *N*-acetyl [U-^{13}C] glucosamine formation was confirmed by ^{1}H NMR (*Figure 4.1*).

Base Epimerisation of N-Acetyl [U-^{13}C] Glucosamine to N-Acetyl [U-^{13}C] Mannosamine (Scheme 2)

[U-^{13}C] GlcNAc (~ 50 mg) was rotary evaporated to dryness and resuspended in 1 mL of water. 100 µL was transferred to a 5 mL glass tube and diluted to 1 mL. The solution was adjusted to pH 11 with 2 M NaOH and stirred at room temperature for 3 days. The resulting solution was neutralised with concentrated HCl. The [U-^{13}C] GlcNAc/ManNAc mixture was desalted on a mixed bed column and purified from any remaining non-carbohydrate material on a Bio-gel P4 column.

Synthesis of N-Acetyl [U-^{13}-C] Neuraminic Acid (Scheme 2)

Chemically synthesised [U-^{13}C] ManNAc (~300 mg, Dr. M. A. Probert, *Figure 1.2*), and [U-^{13}C] sodium pyruvate (CK products Ltd. and Martek Biosciences) (1.1 g) were dissolved in 3.00 mL of 20 mM potassium phosphate buffer pH 7.5, and filtered through a 2 micron filter. Neuraminic acid aldolase (E.C 4.1.3.3)(2 U) was added and the solution was left stirring at 37 °C. A 500 µL aliquot of the reaction mixture was

transferred into a 5 mm NMR tube containing 200 μL D_2O. The reaction was followed by ^{13}C NMR.

[U-^{13}C] Neu5Ac was purified using ion-exchange chromatography (AG 1 X8, 200 mesh, formate form) eluting with 2 M formic acid. The presence of carbohydrate was tested by the sulfuric orcinol method (Lee, 1972). The proton and carbon shifts of [U-^{13}C] Neu5Ac synthesised matched those of a commercial unlabelled sample (Aldrich) (*Figure 4.3*).

Multienzymatic Synthesis of N-Acetyl [U-^{13}C] Lactosamine, (Galβ1-4GlcNAc) (Scheme 4)

[U-^{13}C] galactose (75 mg, 40 % D-isomer, Martek Biosciences), $MgCl_2$ (11.2 mg), ATP (139.2 mg) and UDP-glucose (157.5 mg) were dissolved in 10 mL of 50 mM TRIS-HCl buffer, pH 7.4. The pH was then readjusted to 7.4 with 1 M NaOH. Galactokinase (5 U) (E.C. 2.7.1.6), and galactose-1-phosphate uridyltransferase (5 U) (E.C. 2.7.7.12) were added to the reaction mixture. The initial stages of the reaction were monitored by ^{31}P NMR. The reaction was completed by the addition of [U-^{13}C] GlcNAc (~30 mg), 5 mM $MnCl_2$, and galactosyltransferase (8U) (lactose synthase, E.C. 2.4.1.22). The pH of the reaction mixture was maintained by periodic addition of 0.2 M NaOH. The mixture was incubated for 18 hours and was then desalted (Chelex 100 over Dowex AG50 X12 over Dowex AG3 X4) and purified on Bio-gel P2. Synthesis of [U-^{13}C] *N*-acetyllactosamine was confirmed by ^{1}H NMR (*Figure 4.4*).

Multienzymatic Synthesis of [U-^{13}C] Lactose (Scheme 4)

[U-^{13}C] Lactose was made in an analogous way to [U-^{13}C]-LacNAc except LacNAc was replaced with [U-^{13}C] glucose (~35 mg) (Martek Biosciences) and α-lactalbumin (0.25 mg mL^{-1}). Synthesis of [U-^{13}C] lactose was confirmed by ^{1}H NMR (*Figure 1.5*).

Enzymatic Synthesis of α2,3 Sialyllactose/LacNAc (Scheme 5)

[U-^{13}C] Lactose (8 mg) or [U-^{13}C] LacNAc (10 mg), and *p*-NP [U-^{13}C]-Neu5Ac (~5 mg) was dissolved in 200 μL of 50 mM TRIS-HCl pH 7.4 containing 0.05% NaN_3. The pH of the reaction mixture was maintained at 7.4 by the addition 0.2 M HCl. A 50 μL aliquot of *trans* sialidase enzyme (*E. coli* lysate (Harrison, J. *et al.* 1997)) was added. *p*-NP [U-^{13}C]-Neu5Ac was found to be light-sensitive, so the reaction was incubated in a light-tight container at room temperature and monitored by TLC (30 cm plate, pyridine/butan-1-ol/H_2O/ethanol/acetic acid, 1:1:3:10:0.3). Production of oligo-G_{M3} (R_f 1.69) or α2,3 sialyl *N*-acetyllactosamine (R_f 1.70) could be seen after 2 hours. Additional aliquots of *p*-NP-[U-^{13}C]-Neu5Ac (~5 mg) were dissolved in 50 μL buffer and the pH was adjusted to 7.4 by step-wise addition of 0.2 M HCl and transferred to the reaction mixture, followed by the addition of 50 μL of *E. coli* lysate. This procedure was repeated in 12 hour intervals until all the lactose (R_f 1.49) or LacNAc (R_f 1.50) had been consumed (typically after 60 hours).

Purification was achieved by passing the reaction mixture through a Bio-gel P2 column. NH_4HCO_3 was removed by repeated freeze-drying. The

samples were obtained as the ammonium salt. α2,3 sialyl lactose was obtained in 80% yield, α2,3 sialyl *N*-acetyl lactosamine was obtained in 75% yield. The composition of each sample was confirmed by ^{1}H NMR (*Figures 4.6 & 4.7*) and mass spectrometry (data not shown).

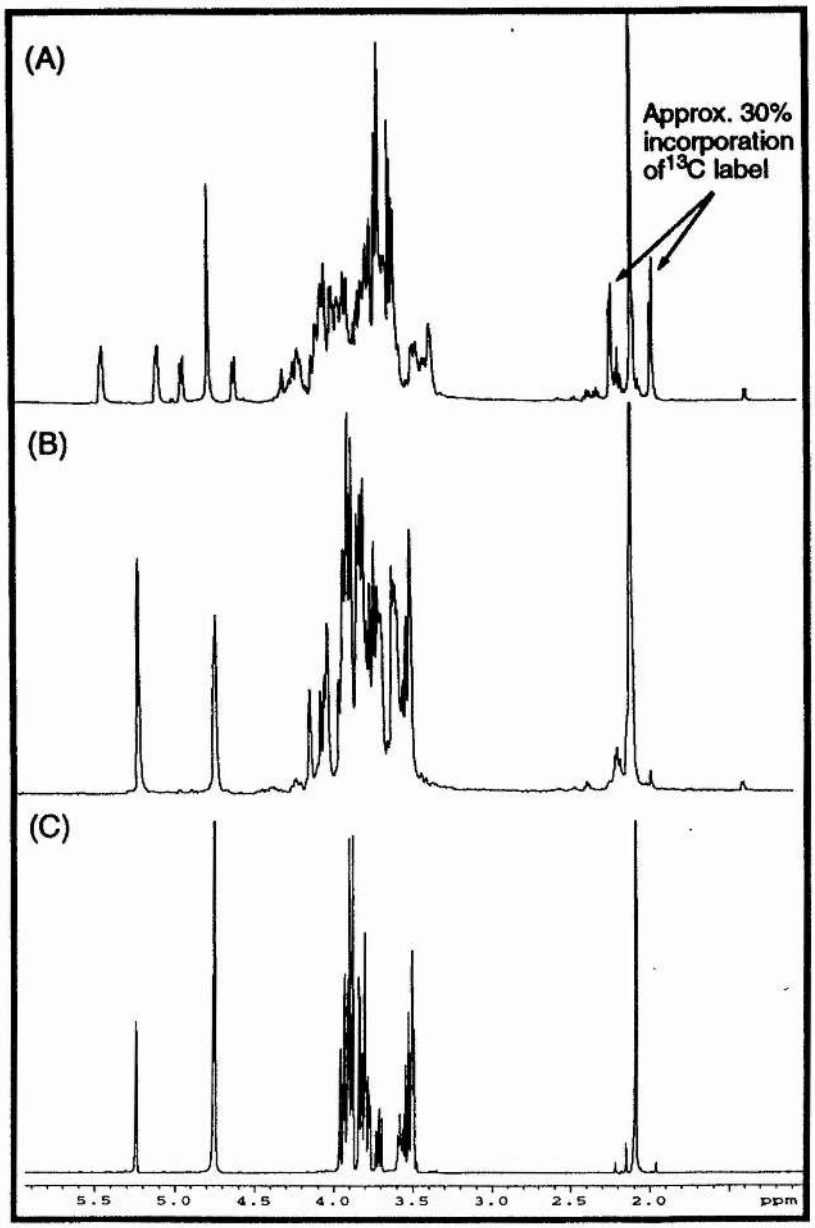


Figure 4.1(A): ^{1}H spectrum of partially purified *N*-acetyl [U-^{13}C] glucosamine, the *N*-acetyl group is approximately 30% ^{13}C enriched. (B): ^{1}H spectrum with ^{13}C GARP decoupling to remove ^{1}H-^{13}C couplings. (C): ^{1}H spectrum of a standard sample (Sigma).

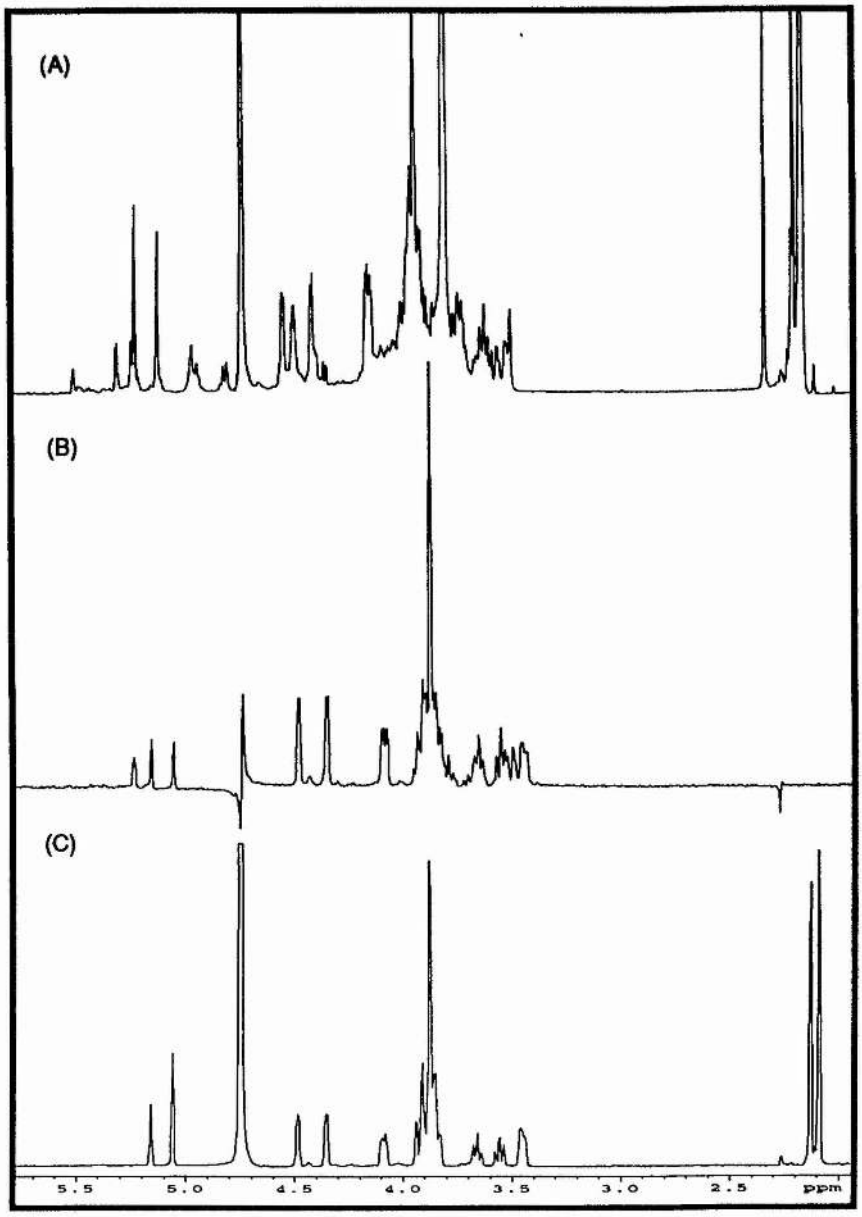


Figure 4.2(A): ^{1}H spectrum of [U-^{13}C] *N*-acetyl mannosamine. (B): 1D ^{13}C-^{1}H HSQC. (C): ^{1}H spectrum of a standard sample (Sigma).

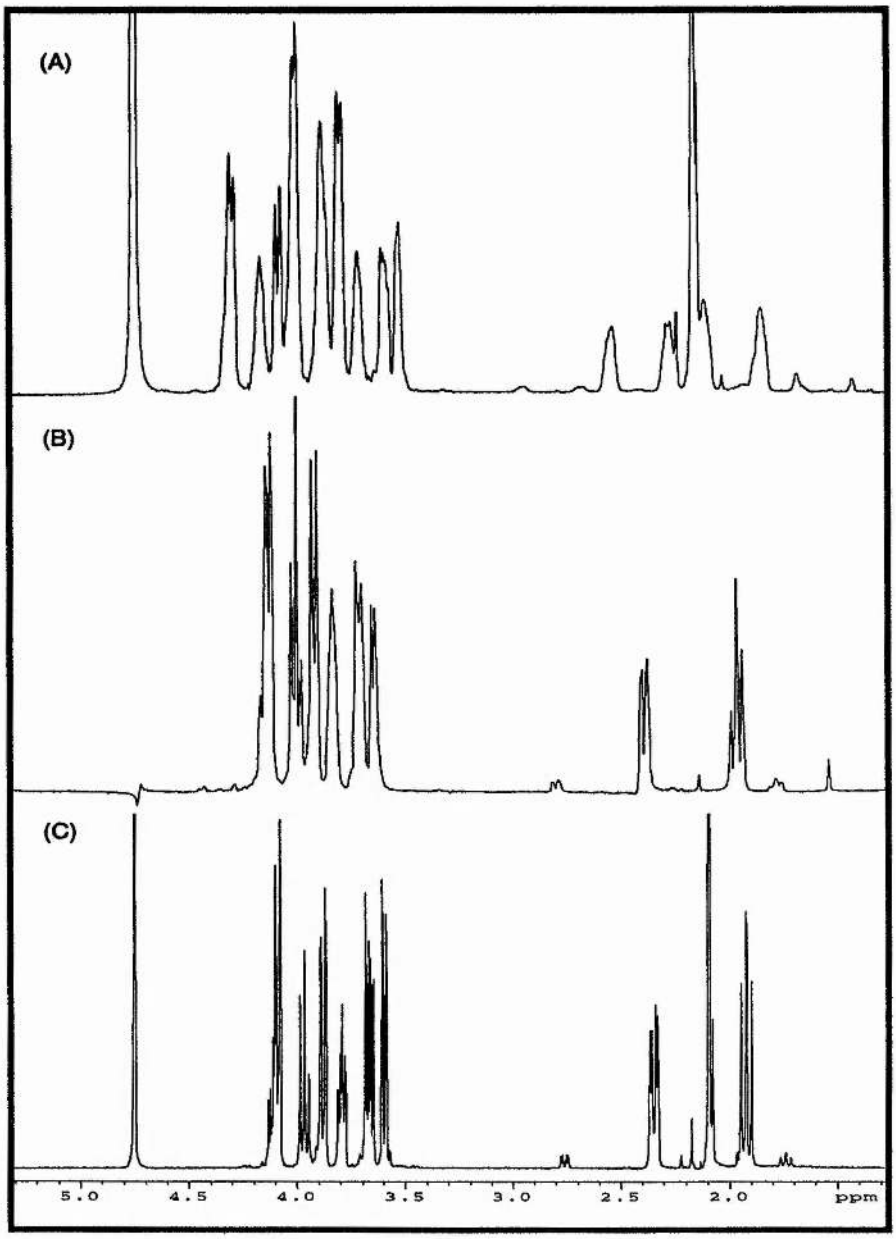


Figure 4.3(A): ^{1}H spectrum of *N*-acetyl [U-^{13}C] neuraminic acid. (B): 1D ^{13}C-^{1}H HSQC. (C): ^{1}H spectrum of a standard sample (Glaxo).

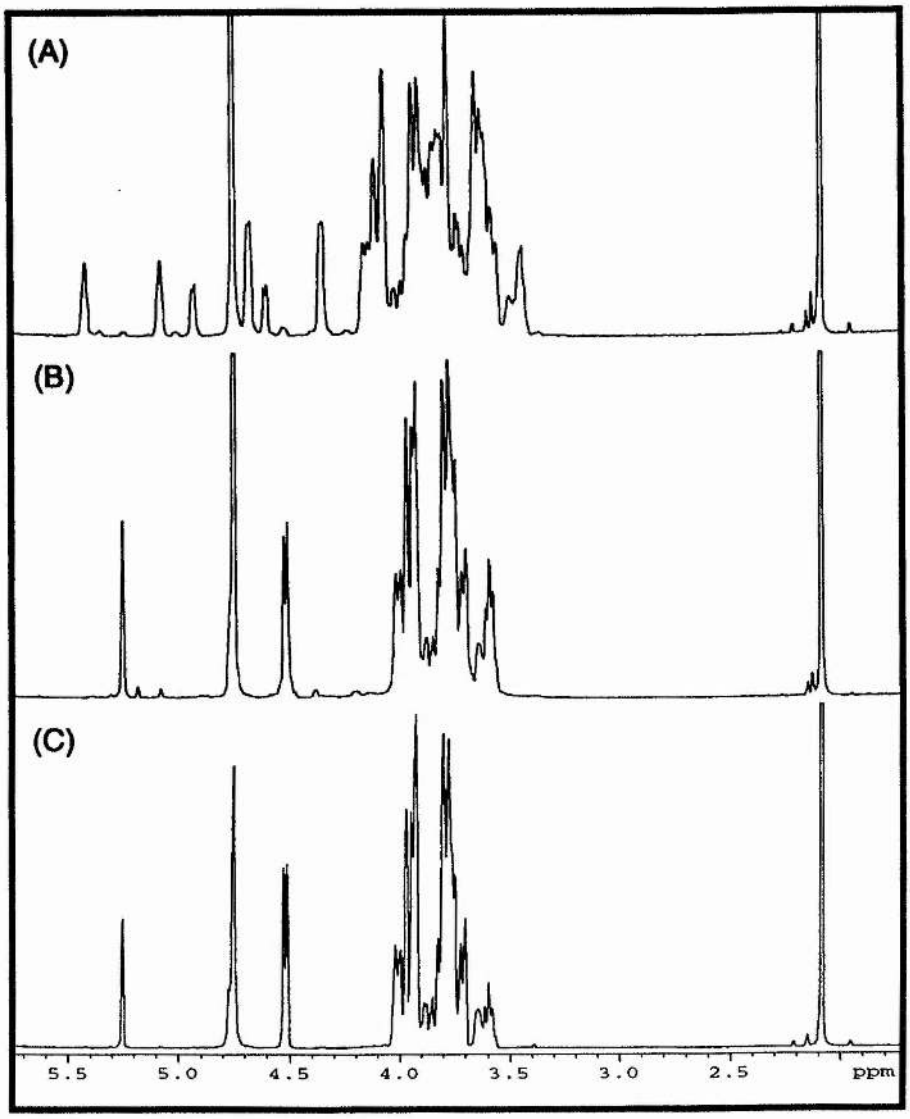


Figure 4.4(A): ^{1}H spectrum of *N*-acetyl [U-^{13}C] lactosamine. (B): ^{1}H spectrum with ^{13}C GARP decoupling to remove ^{1}H-^{13}C couplings. (C): ^{1}H spectrum of a standard sample (Sigma).

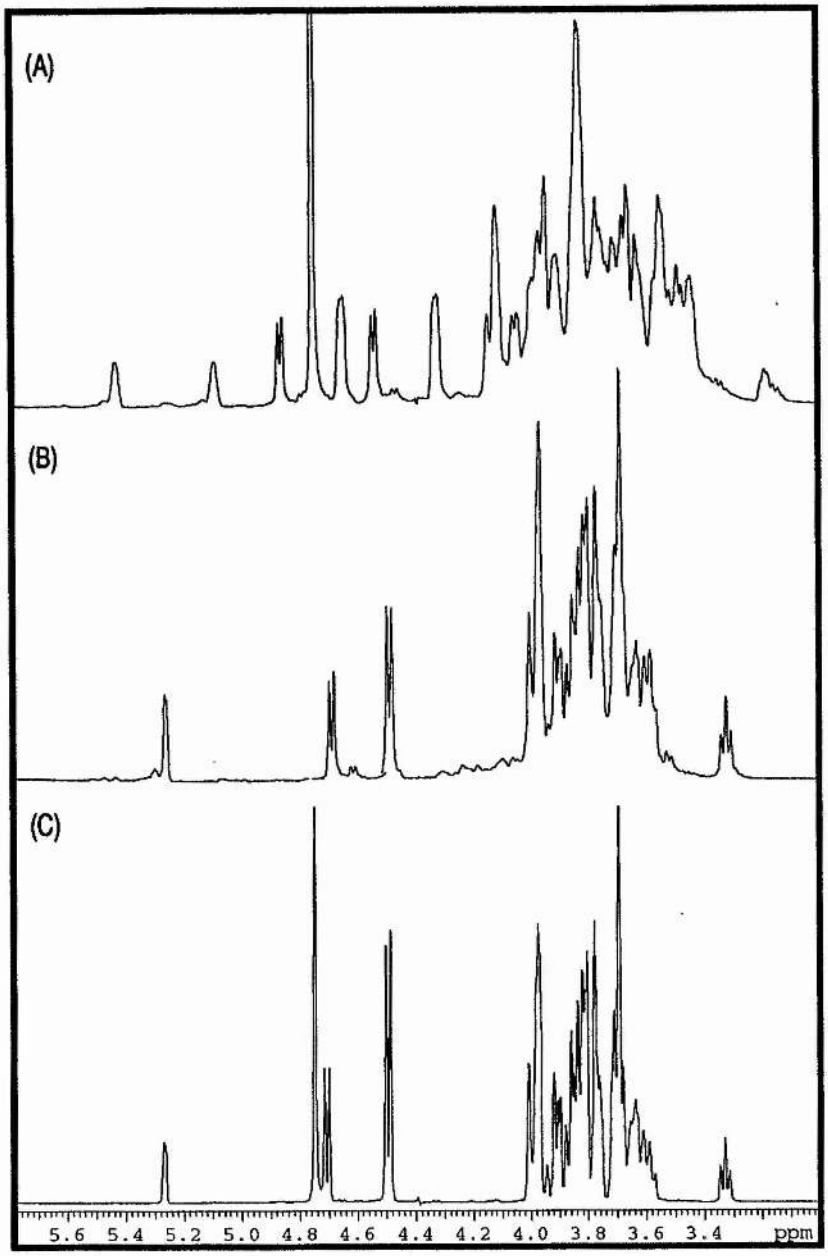


Figure 4.5(A): ^{1}H spectrum of [U-^{13}C] lactose. (B): ^{1}H spectrum with ^{13}C GARP decoupling to remove ^{1}H-^{13}C couplings. (C): ^{1}H spectrum a standard sample (Aldrich).

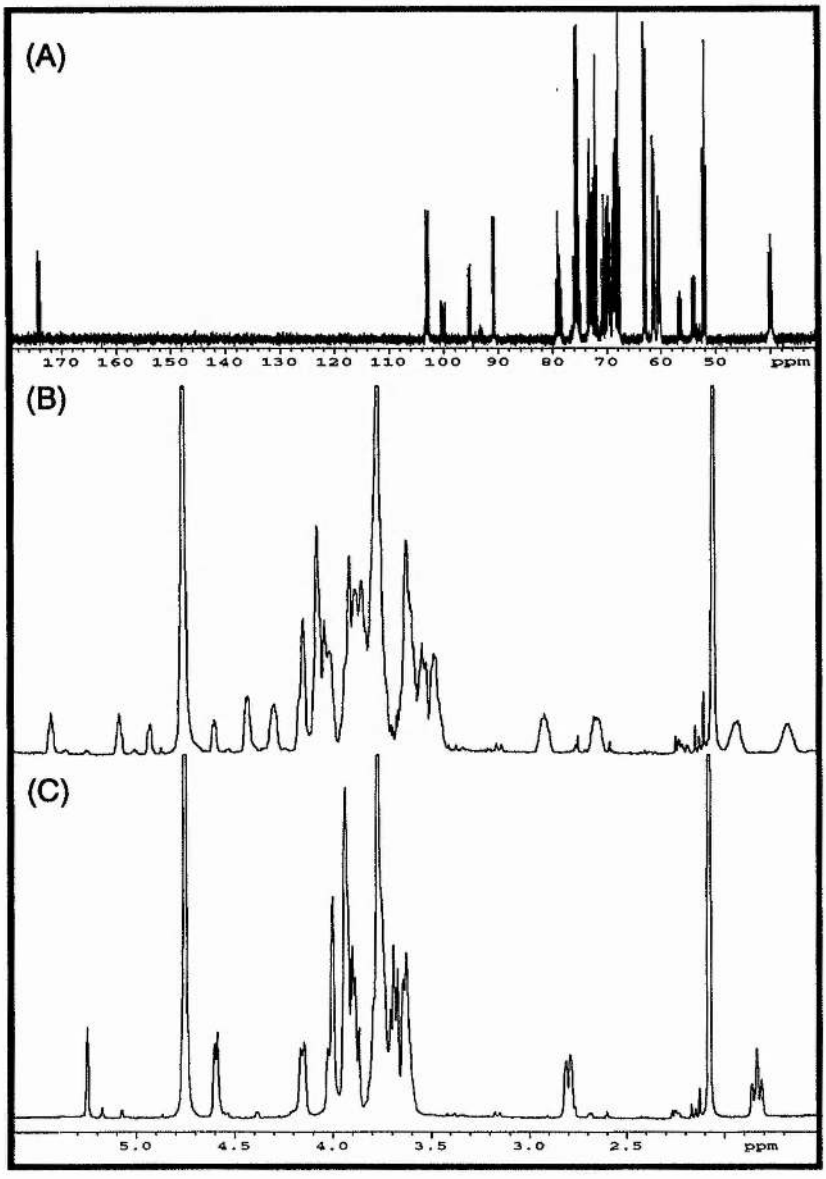


Figure 4.6(A): ^{13}C spectrum of [U-^{13}C] α2,3 sialyl *N*-acetyl lactosamine. (B): ^{1}H spectrum. (C): ^{1}H spectrum with ^{13}C GARP decoupling to remove ^{1}H-^{13}C couplings.

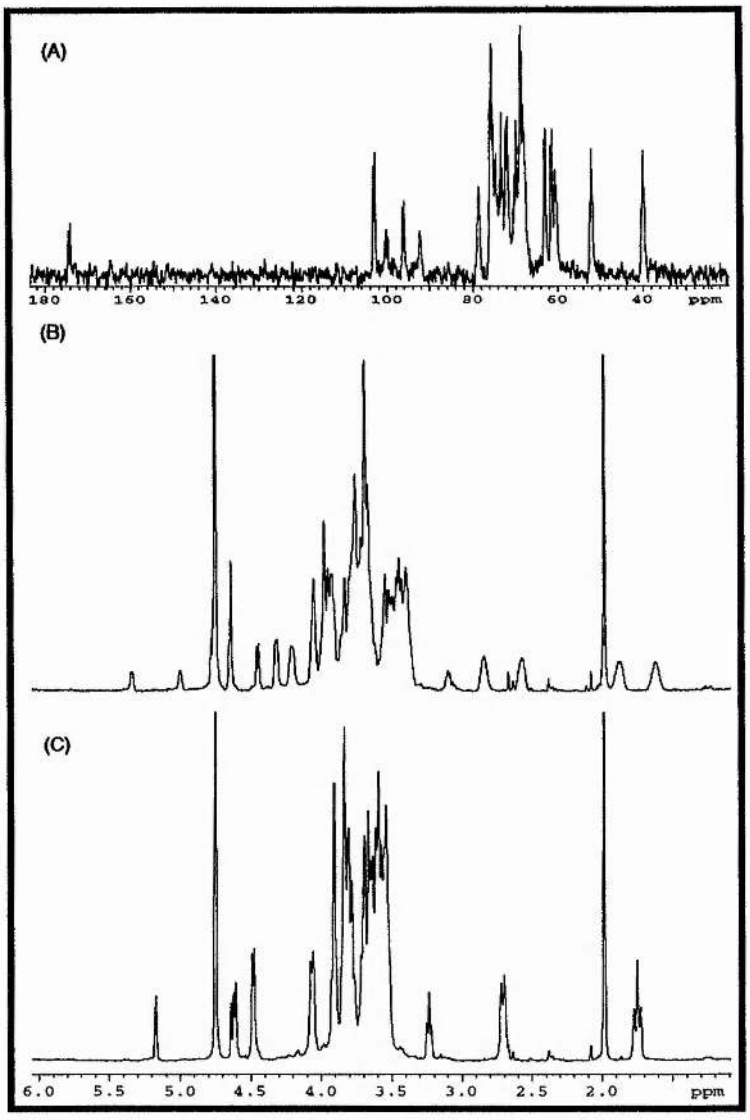


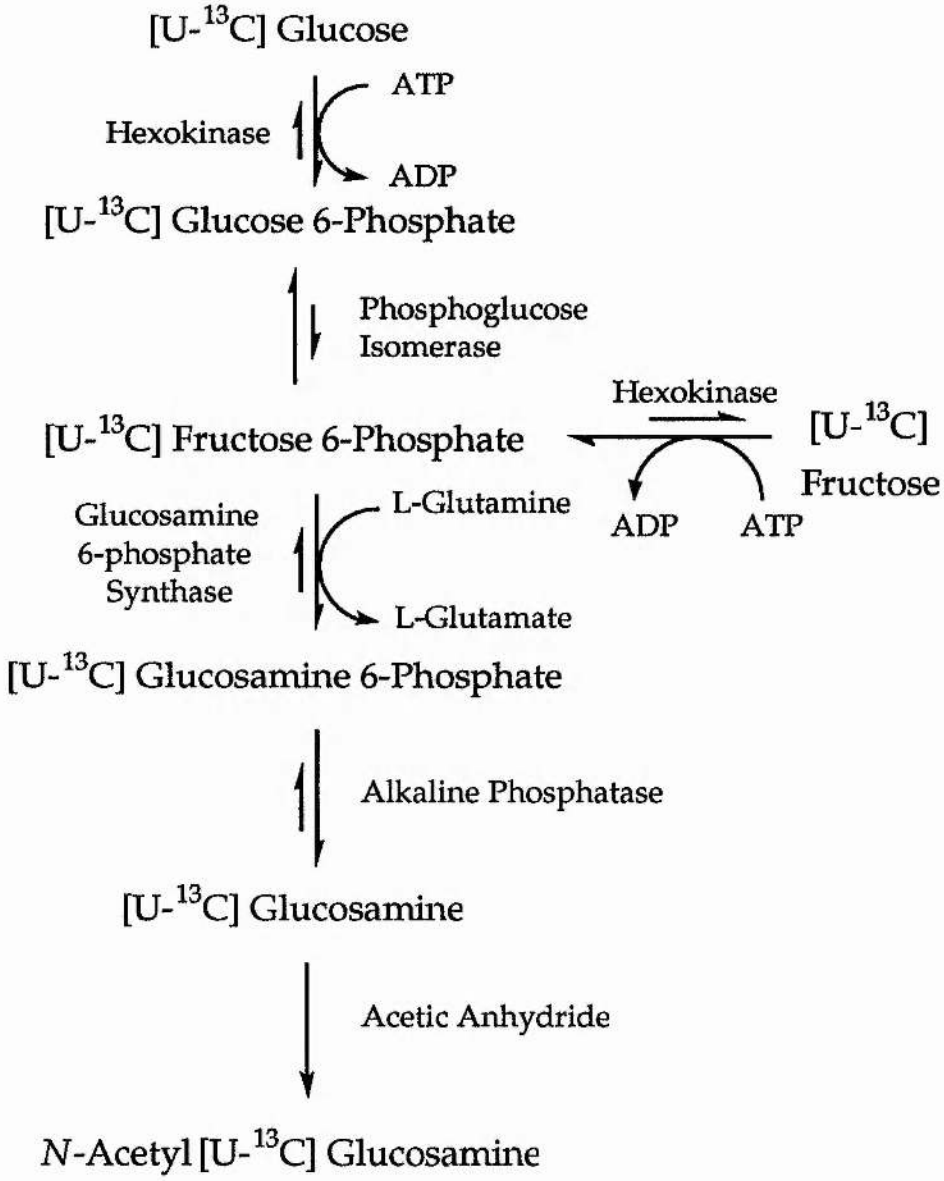
Figure 4.7(A): ^{13}C spectrum of [U-^{13}C] oligo-G_{M3} (α2,3 sialyllactose). (B): ^{1}H spectrum. (C): ^{1}H spectrum with ^{13}C GARP decoupling to remove ^{1}H-^{13}C couplings.

5.3 Results and Discussion

Chemoenzymatic Synthesis of N-Acetyl [U-^{13}C] Mannosamine

A chemoenzymatic synthesis of Neu5Ac from *N*-acetyl glucosamine (GlcNAc) has been published (Simon *et al.* 1988). However, GlcNAc to our knowledge, is not commercially available uniformly carbon-13 labelled. However, [U-^{13}C] GlcNAc may be chemoenzymatically synthesised via a one-pot synthesis of [U-^{13}C] labelled glucosamine 6-phosphate utilising commercially available [U-^{13}C] glucose or fructose, in the biosynthetic pathway (Milton and Homans. 1995; Harris *et al.* 1996) described in scheme 1.

The main disadvantage of this methodology was the unfavourable equilibrium present in the phosphoglucose isomerase (E.C. 5.3.1.9) step. This was overcome by coupling the reaction with glucosamine synthase (L-glutamine: D-fructose 6-phosphate amidotransferase, E.C. 2.1.1.16). Yields were typically in the region of 65% utilising this pathway. Higher yields (>80%) were obtained utilising [U-^{13}C] fructose as a starting material. An additional disadvantage with this scheme is the difficulty in obtaining glucosamine synthase, although, this enzyme may be readily extracted from *E.coli* using the method described by Ghosh and Roseman (1962). More recently glucosamine synthase, has been successfully cloned from *E. coli* and overexpressed in *E. coli* K12 and (Dutka-Malen *et al.* 1988) and yeast (Watzele and Tanner 1989). The protein has also been purified to homogeneity (Badet *et al.* 1987).

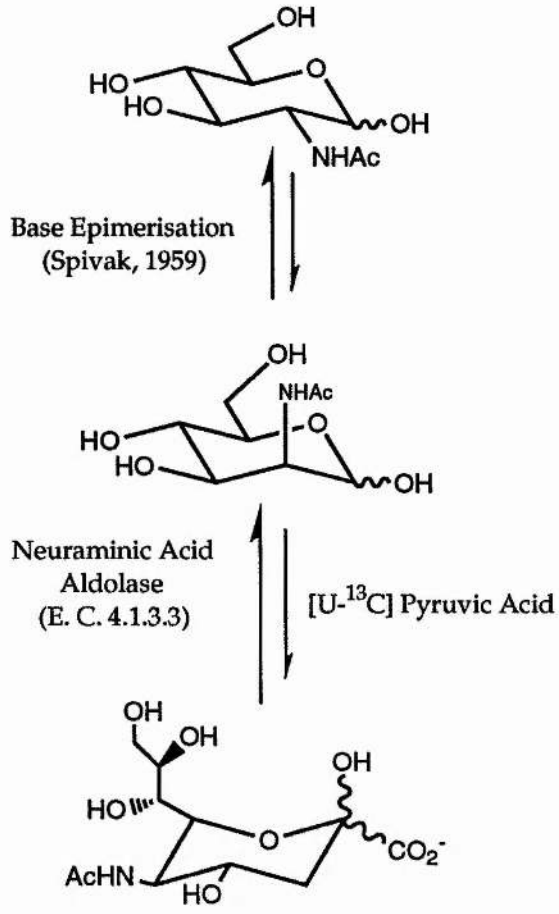


Scheme 1: Chemoenzymatic Synthesis of *N*-Acetyl [U-^{13}C] Glucosamine.

It was the kind donation of a sample of the glucosamine 6-phosphate synthase enzyme from Prof. B. Badet (ENSCP, Paris) which made the chemoenzymatic synthesis of [U-^{13}C] glucosamine possible.

Following de-phosphorylation with alkaline phosphatase (E.C. 3.1.3.1), and *N*-acetylation, the GlcNAc mixture was purified. Initial attempts at purification involved loading the crude glucosamine reaction mixture onto a Dowex AG50W X-12 H^+ and eluting with 2 M acetic acid. However, this resulted in some product decomposition. Acetylation of the crude reaction mixture, desalting and purification by gel-filtration resulted in an almost pure product (*See Figure 4.1*).

A major drawback with this strategy was that it is impossible to selectively *N*-acetylate [U-^{13}C] glucosamine with [U-^{13}C] acetyl chloride (Hare *et al.* 1994) in the crude reaction mixture and a maximum of 30% ^{13}C enrichment was achieved at the *N*-Acetyl position (*See Figure 4.1*). *N*-acetyl glucosamine-2 epimerase has been coupled with Neu5Ac aldolase to synthesis Neu5Ac very efficiently. However, GlcNAc 2-epimerase is no-longer commercially available so the less attractive method of base-epimerisation was employed (Spivak and Roseman 1959). The equilibrium for the epimerisation reaction (*Scheme 2*) lies 3:1 in favour of GlcNAc. The published methods for the separation of ManNAc from GlcNAc using fractional crystallisation (Simon *et al.* 1988) have proven to be difficult to carry-out on a milligram scale. Ion-exchange was also attempted using a Dowex 50W X-2 or Bio-gel P2 columns with borate buffer as eluant (Lee, 1972; Roden *et al.* 1993), however, we were unable to reproduce the results published in the literature.

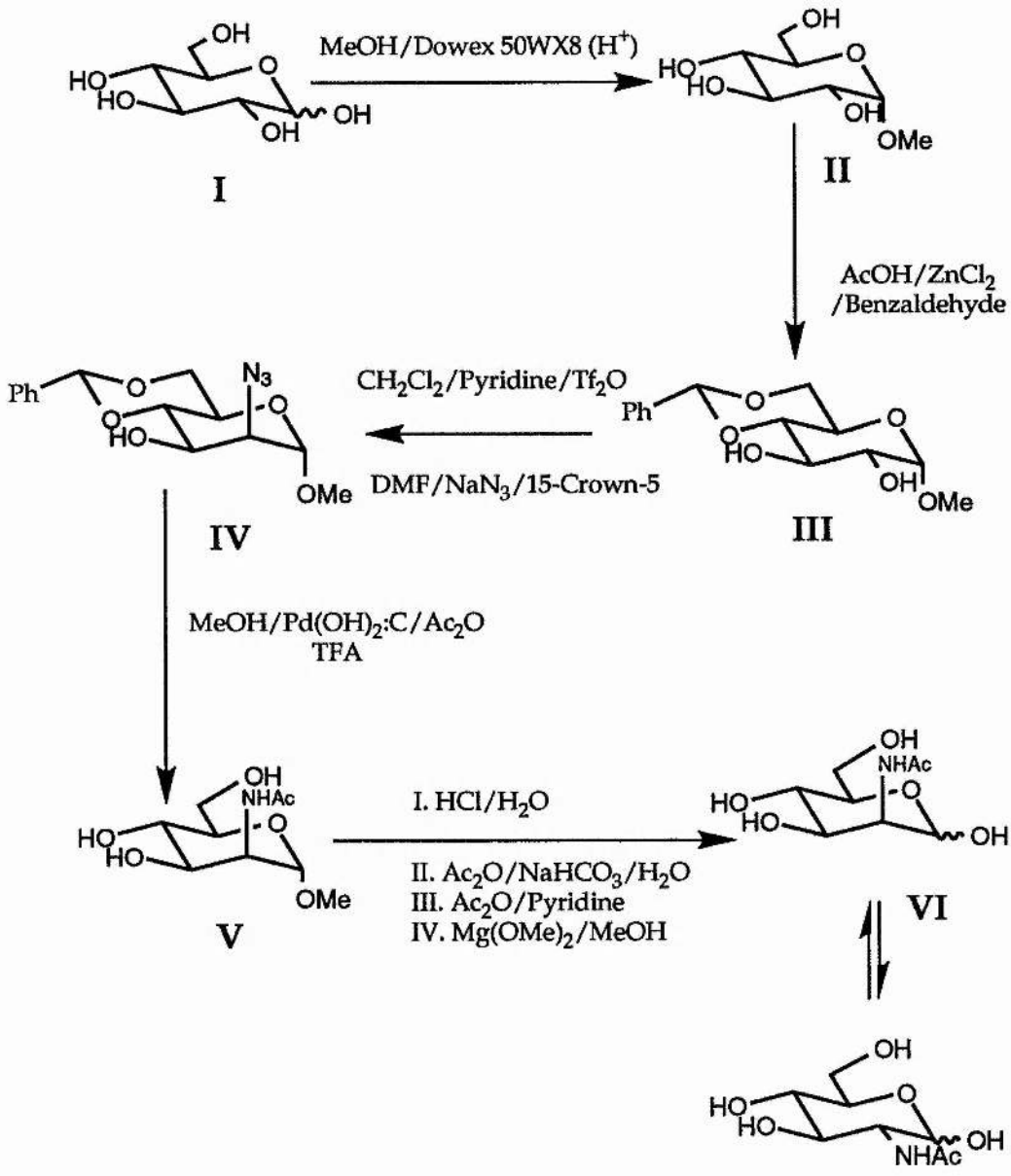


Scheme 2: Base epimerisation of *N*-acetyl glucosamine to *N*-acetyl mannosamine followed by the enzymatic conversion to *N*-acetyl neuraminic acid.

Chemical Synthesis of N-Acetyl [U-^{13}C] Mannosamine

The primary drawback in the chemoenzymatic synthesis of [U-^{13}C] Neu5Ac is the poor yield and difficulties encountered in the purification of ManNAc from the base-epimerisation step (*Scheme* 2). In order to circumvent these problems an alternative chemical synthesis of [U-^{13}C] ManNAc (Scheme 3, performed by Dr. M. A. Probert) was employed (Probert. M. A. *et al.* 1997). Briefly, in the first step, [U-^{13}C] glucose (**I**) is selectively converted to the methyl glycoside with Dowex AG50-X8 (H^+)

and methanol, to give a mixture of α and β anomers. The anomers were separated by fractional crystallisation and the α anomer (**II**) was used in subsequent steps with the β anomer deprotected and recycled. The 4 and 6 hydroxyl positions were selectively protected by reacting **II** with benzaldehyde. The 2 position of **III** was activated by triflation (promoted by the favourable position of the 1-*O*-methyl group) followed by S_N2 displacement of the activated 2-OTf group with N_3^- to yield **IV**. The N_3 group was reduced to NH_2 by hydrogenation over a palladium catalyst followed by *in situ* *N*-acetylation by acetic anhydride. The benzylidene protecting group was only slowly removed by Pd hydrogenation, so a catalytic amount of TFA was added to promote the removal of this group. The methyl glycoside of methyl *N*-acetyl mannosaminide (**V**) was hydrolysed with aqueous acid, although this also results in partial removal of the *N*-acetyl group. The 2-position was re-*N*-acetylated and the free hydroxyls were *O*-acetylated to aid purification by Sephadex G-10. The compound was then de-*O*-acetylated with $Mg(OMe)_2$ to yield [U-^{13}C] ManNAc (**VI**). $Mg(OMe)_2$ is a basic salt and as **VI** is released, the alkali conditions, if left long enough would result in base epimerisation to [U-^{13}C] GlcNAc (Spivak and Roseman 1959), thus the final reaction had to be monitored closely to ensure that the solution could be neutralised as soon as all of the target compound had been deprotected. [U-^{13}C] ManNAc (confirmed by ^{1}H NMR *Figure 4.2*) was purified by gel-filtration. To by-pass the need for basic deprotecting conditions, alternative protecting groups are currently under investigation and will discussed in (Probert *et al.* 1997).



Scheme 3: Chemical synthesis of *N*-acetyl [U-^{13}C] mannosamine.

Enzymatic Synthesis of N-Acetyl [U-^{13}C] Neuraminic Acid

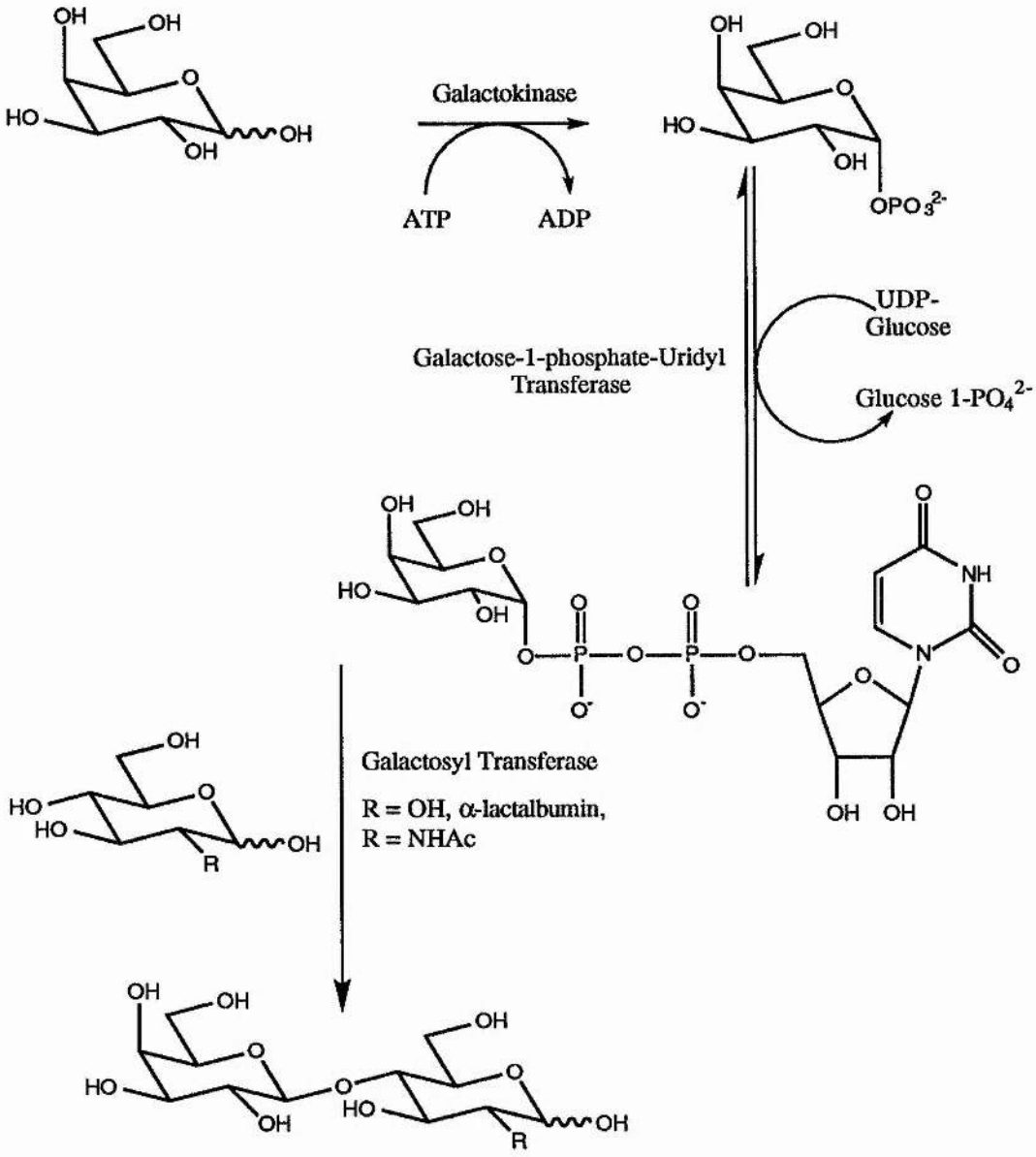
A simplified version of Whitesides' (Simon *et al.* 1988) method for the production of CMP-Neu5Ac was employed to synthesise [U-^{13}C] Neu5Ac (*See Scheme* 2). The equilibrium for this reaction lies strongly in the

direction of the reactants, so a large excess (x8) of [U-^{13}C] pyruvic acid was required to force the reaction to completion. Unreacted ManNAc/GlcNAc was easily separated from Neu5Ac by ion-exchange chromatography, however, unreacted [U-^{13}C] pyruvic acid could not be eluted off the column despite washing the column with high concentrations of acid (up to 4M HCOOH & HCl). This loss of [U-^{13}C] pyruvic acid is not a prohibitive problem when synthesising only a few milligrams of [U-^{13}C] Neu5Ac. If larger quantities of [U-^{13}C] Neu5Ac (> 1 g) are required, alternative reaction conditions would be essential. Kragl (1990, 1992 & 1995) has synthesised Neu5Ac in an enzyme membrane reactor (EMR) using high concentrations (up to 1 M ManNAc) coupled with only a 1.5 times excess of pyruvate. These conditions may be applicable for the large scale enzymatic synthesis of [U-^{13}C] Neu5Ac.

Enzymatic Synthesis of N-Acetyl [U-^{13}C] Lactosamine and [U-^{13}C] Lactose

As discussed in the introduction, sialic acid is often found α2-3/6 linked to *N*-acetyl lactosamine or lactose. A number of different chemical and enzymatic routes have been proposed for the synthesis of galactose β1-4 glucose and its derivatives (Heidlas *et al.* 1992; Ichikawa, 1992; Herrmann *et al.* 1993; Ito and Paulson, 1993; Baisch, G. *et al.* 1996). Published methods for the synthesis of LacNAc, for example, require nucleotide-sugar donor (UDP-galactose) and acceptor (GlcNAc) which are not available carbon-13 labelled. However, by employing the method first developed by Heidlas and co-workers (1992a) and modified by Homans and co-workers (Gilhespy-Muskett, *et al.* 1994), UDP-[U-^{13}C]-Gal was synthesised very efficiently (*Scheme* 3). [U-

^{13}C] GlcNAc has been synthesised using the method described in Scheme 1.



Scheme 4: Enzymatic synthesis of *N*-acetyl lactosamine and lactose.

In the first reaction (see *Scheme* 4), catalysed by galactokinase, [U-^{13}C] galactose and ATP are converted to galactose 1-phosphate and ADP. The equilibrium for this reaction lies strongly in the direction of products.

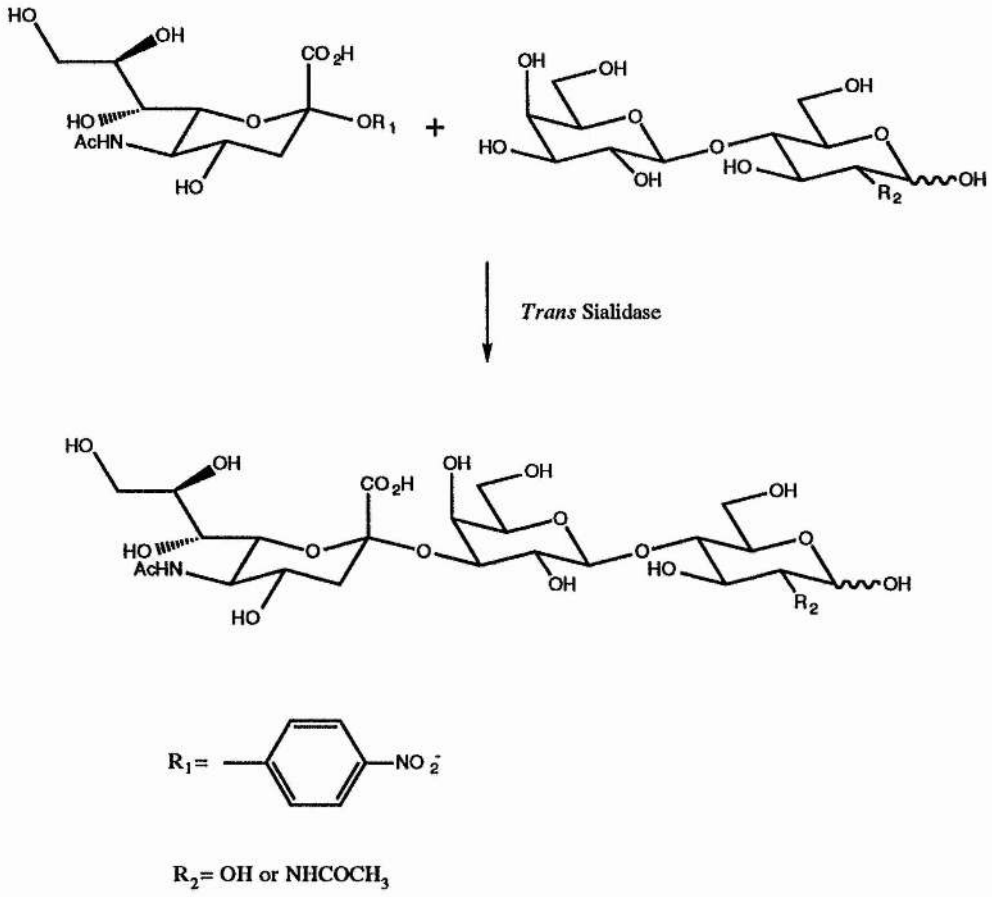
The galactose 1-phosphate product is then converted to the nucleotide-sugar UDP-galactose by transfer of the UDP moiety from UDP-glucose in a reaction catalysed by galactose-1-phosphate uridyl transferase. The formation of UDP-galactose in a yield of ~50 % is obtained from this reaction, at equilibrium. The position of this equilibrium is, however, forced in the direction of products because UDP-galactose is utilised in the final reaction catalysed by lactose synthase. The latter reaction involves the transfer of [U-^{13}C]-galactose to the C-4 position of [U-^{13}C] GlcNAc acceptor. Lactose was made in an analogous way except the presence of a regulatory protein α-lactalbumin was required for the reaction to proceed quickly and to completion. Purification of *N*-acetyl lactosamine and lactose was originally performed on a Bio-gel P4 column. However, in the case of *N*-acetyl lactosamine the relatively large size of the GlcNAc moiety (approx. 1.5 glucose units) resulted in only a partial separation of LacNAc from unreacted GlcNAc. Purification was enhanced by using Bio-gel P2 gel filtration resin.

Enzymatic Routes to Sialylated Oligosaccharides

Sialylated oligosaccharides may be synthesised enzymatically using sialyltransferases which catalyse the transfer of sialic acid from CMP-Neu5Ac to an acceptor (Ichikawa *et al.* 1992; Kittelmann *et al.* 1992; Lin-Chun Liu *et al.* 1992). The major drawback with the use of sialyltransferases is that the procedure is based on stoichiometric reaction of a sugar nucleotide and a mono- or oligosaccharide acceptor, this requires a separate preparation of the sugar nucleotide and suffers from product inhibition caused by the released nucleotide mono phosphate. This problem was tackled by Wong and co-workers (Ichikawa. *et al.* 1992;

Kittelmann *et al.* 1992; Lin-Chun Liu *et al.* 1992), who developed a very efficient *in situ* regeneration system. α2,3 sialyl *N*-acetyllactosamine has been synthesised using the *in situ* regeneration scheme in very poor yield (<20 %). However, α2,3 sialyltransferase is not readily available. The insolubility of the protein also makes it difficult to handle in synthetic reactions. Baisch and co-workers (Baisch *et al.* 1996a,b) have recently published the expression and purification of a truncated soluble α2,3 sialyltransferase in insect cells. However, an additional disadvantage with α2,3 sialyltransferase is that it will only accept LacNAc or acyl derivatives of LacNAc as substrates. Therefore an alternative synthetic scheme would be required for the synthesis of oligo-G_{M3}. Trypomastigotes of *Trypanosoma cruzi* have been found to express a developmentally regulated enzyme, unique to *T. cruzi*, α2,3-*trans*-sialidase. *Trans*-sialidases are unusual in that they transfer Neu5Ac from an α2,3-sialylated oligosaccharide and not CMP-Neu5Ac, the normal donor substrate for mammalian sialyltransferases. This ability of the *trans*-sialidase to utilise α2,3-sialylated sugars as donor substrates prompted experiments to investigate the potential of the enzyme to transfer Neu5Ac from synthetic α–sialosides such as 2-*O*-(*p*-Nitrophenyl)-*N*-acetyl-neuraminic acid (*p*-NP-Neu5Ac), Neu5Acα2-3Galβ1-*p*-NP, and 2′-(4-methylumbelliferyl)-*N*-acetyl-α-neuraminic acid (Scudder *et al.* 1993). *Trans*-sialidase has been expressed in *E. coli* (Harrison *et al.* 1997) and the crude cell lysate was active enough for use in enzymatic synthesis.

Despite the fact that *p*-NP-Neu5Ac was shown to be a rather kinetically poor substrate for *trans*-sialidase, it is easily chemically synthesised in high yield and the substrate is active enough for the the synthesis of [U-^{13}C] oligo-G_{M3} (α2,3 sialyllactose) and [U-^{13}C] α2,3 sialyl *N*-acetyl lactoseamine on the milligram scale in high yield (> 75%). *p*-NP-[U-^{13}C] Neu5Ac was synthesised by Dr. M. A. Probert.



Scheme 5: Enzymatic synthesis of α2,3 sialyllactose and α2,3 sialyl *N*-acetyl lactosamine.

4.4 Conclusions

Uniformly carbon-13 labelled Neu5Ac has been synthesised utilising a predominately enzymatic route to the intermediate compound ManNAc. However, the relative difficulty in obtaining glucosamine synthase and difficulties in purifying ManNAc from the base-epimerisation step made the alternative chemical synthesis of ManNAc more attractive despite this route requiring more time and materials. GlcNAc was obtained via both routes either as a product in scheme 1 or as a by-product from the last de-protection step in scheme 3. GlcNAc and glucose have been coupled to galactose in high yield via a one pot enzymatic synthesis described in scheme 4. Oligo-G_{M3} and α2-3 sialyl *N*-acetyl lactosamine were produced by employing a novel enzymatic synthetic route which uses an enzyme (*trans*-sialidase) which does not require the traditional sugar nucleotide, CMP-Neu5Ac, as a donor substrate but accepts a synthetic α-sialoside, *p*-NP-[U-^{13}C]Neu5Ac, as a substrate. Methods described in this chapter have allowed access to milligram quantities of uniformly enriched carbohydrate moieties crucial in the NMR investigations of biologically important systems. The successful synthesis of oligo-G_{M3} and α2-3 sialyl *N*-acetyl lactosamine moieties was a prerequisite for the synthesis of more complex gangliosides and sialyl Lewis^x antigen, respectively.

Chapter 5

Heteronuclear NMR Investigation of the Solution Conformations of Oligo-G_{M3} in Free Solution and in Complex with Triticum vulgaris Lectin

Abstract

Heteronuclear NMR spectroscopy was employed to study the structure and dynamics of a carbon-13 enriched glycan moiety of the ganglioside G_{M3}: oligo-G_{M3} (*N*-acetyl [U-^{13}C] neuraminic acid (α2-3) [U-^{13}C] lactose, Neu5Acα2-3Galβ1-4Glc) in free solution and in complex with *Triticum vulgaris* lectin (wheat germ agglutinin, WGA). Structural information on the free sugar was obtained from three dimensional (3D) ^{13}C-edited ^{1}H-^{1}H ROESY (ROESY-HSQC) data and long range carbon-proton and carbon-carbon coupling constants. 3D ^{13}C-edited ROESY experiments provided unambiguous assignments of all ROESY cross-peaks, and proved that one of the ROE cross-peaks reported in the literature was assigned incorrectly. 3D ROESY-HSQC experiments also revealed two additional ROEs across the Galβ1-4Glc glycosidic linkage. The bound conformation of oligo-G_{M3} in complex with WGA was elucidated using 3D ^{13}C-edited transferred NOESY (TRNOESY-HSQC) experiments. Theoretical TRNOEs for the Neu5Acα2-3Gal linkage back-calculated from the X-ray crystal structures matched those measured experimentally and confirmed that this linkage adopts the ϕ, ψ ~70°, 0° conformation when bound to WGA in solution. Qualitative information on the mobility of complexed oligo-G_{M3} was extracted from ^{13}C -^{1}H HSQC experiments.

5.1 Introduction

The trisaccharide α2,3 sialyllactose (oligo-G_{M3})(Neu5Acα2-3Galβ1-4Glc) (*See Figure 5.1*) forms the oligosaccharide moiety of the ganglioside G_{M3} (Neu5Acα2-3Galβ1-4Glcβ1-ceramide). G_{M3} itself forms the core structure of many biologically important glycoconjugates such as G_{M1} which is the functional receptor for cholera and heat labile enterotoxins. G_{M3} has been shown to modulate cell growth by inhibiting EGF-dependent kinase tyrosine phosphorylation of the EGF-receptor. Myeloma cells are often characterised by increased expression of G_{M3} on the cellular surface, such cells are known to shed a considerable amount of membrane material containing a high concentration of G_{M3}. This appears to generate a negative immune response resulting in the enhancement of melanoma growth. A G_{M3}-antibody interaction has been identified as the major factor responsible for the negative response, thus, there is a great deal of interest in F_{ab}- G_{M3} interactions (Otting and Wuthrich, 1989).

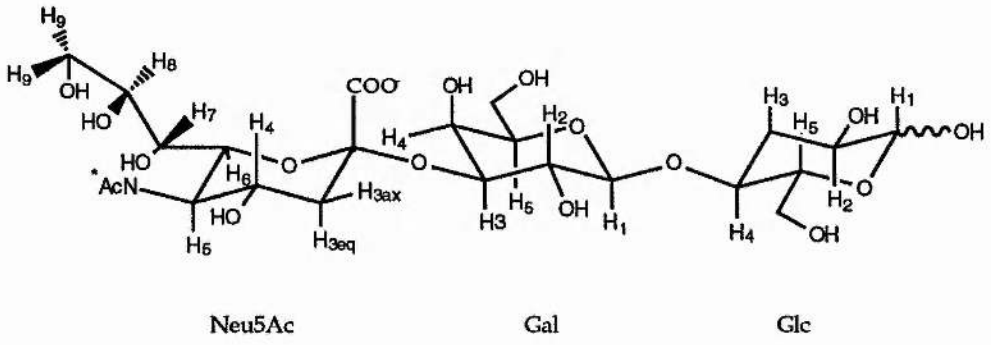


Figure 5.1: Schematic representation of oligo-G_{M3} including the numbering scheme used in the text. In NMR experiments all carbons are carbon-13 enriched (~99 %) except for the *N-acetyl group.

G_{M3} is also a useful model compound for the study of ganglioside - protein interactions. G_{M3} is known to interact with proteins which are readily available and well characterised and may serve as suitable systems for the study of changes in the structure and dynamics of receptor headgroups on protein binding. Wheat germ agglutinin (WGA), a plant lectin, is one such protein. It is known to aggregate vesicles containing a number of derivatives of G_{M3}, and several high resolution structures of WGA - G_{M3} complexes have been determined by X-ray crystallography (Wright, 1980; 1984; 1990; 1992; Wright and Jäger, 1993).

Various high-resolution NMR techniques have been used to determine the three-dimensional structure and dynamics of G_{M3} in a number of chemical environments. Intact G_{M3} fixed in micelles and oligo-G_{M3} in free aqueous solution has been studied by homonuclear proton NMR spectroscopy (Siebert *et al.* 1992). Following the recent developments of high resolution multidimensional heteronuclear NMR experiments (Bax, 1994; Edison *et al.* 1994) and the synthesis of highly carbon-13 enriched oligo-G_{M3} (discussed in the chapter 4), a rigorous investigation of the structure and dynamics of oligo-G_{M3} in free solution and in complex with WGA is now possible, although, a heteronuclear (^{1}H and ^{13}C) NMR study of G_{M3} structure has been reported (Aubin*et al.* 1993): the membrane conformation of *N*-acetyl [1, 2, 3,-^{13}C]-neuraminic acid labelled G_{M3} was studied on the surface of a magnetically orientated membrane by measuring ^{13}C-^{13}C and ^{1}H-^{13}C dipolar splittings in a manner akin to solid state NMR. However, the information obtained from such solid state NMR parameters is very limited. Only the orientation of the Neu5Ac pyranose ring relative to the membrane surface could be extracted from

such studies. The aim of the current investigation was to describe the three-dimensional solution structures and dynamics of oligo-G_{M3} with the application of three dimensional ROE/NOESY-HSQC experiments to assign unambiguously all ROE/NOE cross-peaks and to record additional structural information such as $^3J_{CC}$ and $^3J_{CH}$ coupling constants. The bound conformation of oligo-G_{M3} was also probed by acquiring 3D carbon-edited transferred NOESY experiments.

5.3 Materials and Methods

Free Solution Studies

[U-^{13}C] oligo-G_{M3} . NH_4^+ (~7.5 mg) was dissolved and lyophilised into 99.96% D_2O three times followed by dissolution into 700 µL D_2O.

NMR Experiments

NMR spectra were obtained at 300K with a ^{1}H reference frequency of 500 MHz on a Varian Unity$^+$ spectrometer equipped with a self shielded z gradient triple resonance probe.

Homonuclear ^{13}C (^{1}H decoupled) two dimensional (2D) coupling correlated spectroscopy experiment (COSY) was recorded in magnitude mode with a spectral width of 12 kHz, 512 t_1 increments, 128 scans per increment with 8K complex points in t_2, resulting in a total acquisition time of ~24 hours.

The remaining spectra were recorded in the phase-sensitive mode with

use of the States' (1982) method for quadrature detection. 2D HCCH-COSY experiments were recorded using the three-dimensional (3D) pulse scheme of Bax (1990). A total 1K complex and 2K complex points were acquired in the t_1 and t_2 dimensions respectively, with spectral widths of 2 kHz and 10 kHz. Heteronuclear ^{13}C- ^{1}H correlated single-quantum coherence (HSQC) experiment was acquired with 512 t_1 increments, 32 scans per increment with 4K complex points in t_2. ^{1}H-^{13}C couplings evolving in t_2 were removed by ^{13}C GARP decoupling. Total acquisition time was ~12 hours.

2D (Bax *et al.* 1992) & 3D (Bax *et al.* 1994) gradient-enhanced long range carbon-carbon coupling (LRCC) experiments were recorded with a proton sweep width of 2 kHz, consisting of 1024 complex points and a ^{13}C sweep-width of 11 kHz in t_1 (128 complex points) and a 6.5 kHz (32 complex points) in t_2 for the 3D experiment. 64 and 8 scans were acquired per increment for the 2D and 3D experiments, respectively. ^{13}C-^{13}C couplings evolve and are refocussed during the delay (2T) of 11.1 ms. Total acquisition times for the 2D and 3D experiments were ~12 hours and ~48 hours, respectively. The values of the long range coupling constants are derived from the ratios of cross-peaks obtained in the spectrum in the following manner:

$$^{n}J_{CC} = \frac{1}{2\pi T}\tan^{-1}\left(\sqrt{\frac{V_{Ir}}{V_{ref}}}\right) \qquad \textit{Eqn 5.1}$$

where V_{Ir} is the volume of the long range (H_aC_{a+n}) cross-peak and V_{ref} is the reference cross-peak (H_aC_a) volume.

Trans-glycosidic carbon-proton coupling constants were measured using the 2D constant-time pulse sequence (CT-LRCH) of Bax (1994). Proton and carbon sweep-widths were identical to those used in the 2D LRCC spectrum with 4096 complex points in t_2 and 256 t_1 increments consisting of 16 scans per increment.

Three-dimensional offset compensated ^{13}C-edited ROESY experiments were acquired using a conventional ROESY-HSQC pulse sequence with spectral widths of 2 kHz, 6.5 kHz, and 2 kHz; and 128, 32, and 1024 complex points in t_1, t_2, and t_3 respectively. The proton transmitter offset was placed 500 Hz downfield of the lowest field resonance during the spin-lock period to minimise coherence transfer effects. To optimise digital resolution the Neu5Ac C5 and C3 resonances (54 ppm and 40 ppm, respectively) were folded-in once. The incremental delay at time zero in t_2 was calculated according to (Edison *et al.* 1994) to give a processed spectrum where folded peaks have an opposite amplitude to non-folded peaks when zero and first order phase corrections are set to 90° and -180°, respectively. A total of 128, 32, and 1K complex points were acquired in t_1, t_2, and t_3 respectively. The effective field for spin-locking was 2 kHz and was applied for a mixing time of 250 ms. Prior to Fourier transformation, data were apodised with cosine-bell functions, and the t_1, and t_2, dimensions were zero filled to 256 and 64 complex points, respectively. Total acquisition time for the 3D spectrum was ~55 hours.

Bound State Conformational Studies.

To determine the optimal TRNOE conditions, unlabelled oligo-G_{M3} (Sigma) (2.5 mg) was dissolved in 700 µL of 100 mM deuterated sodium

phosphate buffer (NaPi) pD 6.4 containing 150 mM NaCl and 14 mM NaN_3. Commercially available WGA (Sigma) was purchased as a lyophilised salt free powder and was added, without exchange into D_2O, in 0.5 mg aliquots. 2D phase sensitive NOESY spectra were recorded after each addition of protein. NOESY experiments consisted of 256 t_1 increments with 8 scans per t_1 increment and a receiver delay of 1 s, and 1K complex data points in t_2. An initial mixing time of 100 ms was used. At a ligand:protein ratio of 33:1 (the average molecular weight of the WGA dimer is 36 kDa), the TRNOE intensity vs. line broadening appeared to be optimal.

Homonuclear 2D offset compensated TRROESY experiments were recorded using a conventional ROESY experiment as described previously (Rutherford *et al.* 1994). The rf carrier frequency was placed 500 Hz up field relative to the highest field resonance to minimise coherence transfer effects. The effective field for spin-locking was 2 kHz and was applied for a mixing time of 75 ms. A total of 1024 t_1 increments, 32 scans per increment with 4K data points were accumulated with a spectral width of 3.7 kHz.

All 1D and 2D time domain data were zero filled once, multiplied by a squared cosine bell function prior to Fourier transformation.

[U-^{13}C] oligo-G_{M3} . NH_4^+ (2.5 mg) was dissolved and lyophilised into 99.96% D_2O three times followed by dissolution into deuterated 100 mM NaPi - 150 mM NaCl pD 6.4 buffer (*vide supra*) to which was added WGA (4 mg). The sample was transferred into a 5 mm NMR tube (Wilmad

528pp). 3D-NOESY-HSQC experiments were acquired with NOESY mixing times of 75 ms. The ^{1}H sweep width was 2 kHz in t_1 and t_3 and the ^{13}C sweep-width was set to 6.5 kHz in t_2. As described previously, to optimise digital resolution the Neu5Ac C5 and C3 resonances (54 ppm and 40 ppm, respectively) were folded-in once. A total of 128, 32, and 1K complex points were acquired in t_1, t_2, and t_3 respectively to give a total acquisition time of ~55 hours. Prior to Fourier transformation, data were apodised with cosine-bell functions, and the t_1, and t_2, dimensions were zero filled to 256 and 64 complex points, respectively.

2D NOESY and ROESY peak volumes were measured using the 2D integration routine contained in Varian VNMR software. 3D peak volumes were measured by summing the volume of cross-peaks over all the 2D *F1 F3* planes in which they appear.

Calculation of the Correlation Time (τ_c) for Wheat Germ Agglutinin.

The correlation time for WGA was determined from light scattering studies as follows: A 3 mg mL^{-1} sample of WGA in 0.1 M NaPi - 0.15 M NaCl pH 6.0 was prepared; A 100 µL sample was loaded on to a DYNAPRO-801 dynamic light scattering recorder; R_H, the hydrodynamic radius, was measured to be 0.3 nm ± 0.2. τ_c can then be approximated to the rotational correlation time given by Debye (1929):

$$\tau_c = \frac{4\pi\eta R_H^{\,3}}{3k_bT} \qquad \textit{Eqn 5.1}$$

Where k_b is the Boltzmann constant, η is the solvent viscosity, and T is the absolute temperature. The τ_c value for the WGA dimer was calculated to be ~27 ns. Light scattering studies results were indicative of a monodisperse solution and no aggregation occurred upon addition of oligo-G_{M3}.

Conventions

The torsion angles ϕ(phi), ψ(psi) are analogous to ϕ_H and ψ_H in IUPAC convention, and are defined as H1-C1-O1-CX and C1-O1-CX-HX, where CX and HX are aglyconic atoms. Neu5Ac ϕ torsion angle was defined as C1-C2-O2-CX, torsion angles around the glycerol side chain are defined as ϕ_1: H6-C6-C7-H7, ϕ_2: H7-C7-C8-H8, and ϕ_3: H8-C8-C9-O9.

Molecular Modelling

Computer models were assembled using Biosym's Insight II molecular graphics package (v 2.8, Biosym Technologies, San Diego, CA) running on a Silicon Graphics Indigo 2 Challenge computer. Random structures were generated by dynamical quenching. An initial structure was built with pyranose rings in the $^4C_1$ ($^2C_5$ for Neu5Ac) chair conformation with trial values of phi (ϕ) and psi (ψ), and subjected to 200 ps of unrestrained molecular dynamics at 750K, during which the torsional terms are scaled by a factor of 7 to prevent excessive ring puckering. A random structure was saved every 10 ps. Energy minimisation by restrained simulated annealing was achieved as follows: models were equilibrated for 10 ps with a thermal bath at temperatures 500K, 450K, 350K, 300K, and then successively for 1 ps in decreasing steps of 10K, followed by a further 1 ps

at 5K. The system was minimised using the steepest descents algorithm until the maximum derivative was less than 0.04 kJ $mol^{-1}Å^{-1}$. Restraints listed in table 4 were applied as a biharmonic function (discussed in chapter 1). NOE contacts were arbitrarily assigned as strong (1.8Å - 2.7Å), medium (1.8Å - 3.3Å) or weak (1.8Å - 5.0Å). TRNOE contacts were assigned as strong (1.8Å - 2.7Å) or weak (1.8Å - 5.0 Å).

All dynamics simulations were performed *in vacuo* with a dielectric constant (ε) of 80.0 for simulations describing the motion of the free oligosaccharide. To mimic, in part, the micro-environment of the protein binding site a dielectric constant of 4.0 was chosen for dynamical simulated annealing simulations using TRNOE distance restraints. The AMBER force-field (Singh *et al.*, 1986) with a carbohydrate parameter set developed by Homans (Homans 1990) was used with the exo-anomeric potentials set to zero. Simulations lasted a total of 5.01 ns real time, with the last 5 ns used in further analysis.

Construction of the WGA Binding Site

The three dimensional co-ordinates for WGA isolectin 1 in complex with oligo-G_{M3} and T5-antigen were obtained from the Brookhaven Protein DataBank. The correct structure of the WGA dimer in complex with oligo-G_{M3} was constructed as follows: Out of two symmetrically independent WGA monomers in a unit cell, we have used the coordinates of molecule A which corresponds to protomer 1 in the literature (Wright, 1992). WGA 1 coordinates were loaded into Biosym's Insight II graphics programme and a crystal lattice (comprising of 6 dimers was recreated by copying and rotating the coordinates according to the

space group C_2 symmetry rules. One dimer and its associated ligand was used as a basis for full relaxation matrix calculations (model **I**). Three other bound conformers of oligo-G_{M3} have been solved by X-ray crystallography and these have been used to generate models **II**, **III**, and **IV** by docking these sugars into the binding site generated for model **I**. Model **V** was created by changing the ϕ, ψ dihedral angles for Galβ1-4Glc linkage in model **I** to 40°, 0°, respectively. The correct structure of WGA1 in complex with T5 was generated by copying and rotating the coordinates of the T5 molecules only 26 times according to $P2_12_12$ space group rules. The Galβ1-3R moiety of T5 was replaced by Galβ1-4Glc (ϕ, ψ set to 40° & 0°, respectively) and minimised for 2000 steps of conjugate gradient minimisation to give model **VI** which represents the C binding site.

5.2.1 Full Relaxation Matrix Calculations for an Exchanging System

Theoretical NOEs, ROEs, and TRNOEs were calculated using the in-house software 'MDNOE2'. For TRNOE calculations, the coordinates of all protein non-exchangable protons within 8 Å of the Neu5Ac residue of the models described above were inputted into MDNOE2.

To calculate the development of transferred NOEs theoretically, a full relaxation matrix (**R**) coupled with a kinetic matrix (**K**) describing the chemical exchange is required. For an exchanging system the time development of cross-peak intensities as a function of mixing time is given by

$$\mathbf{M'}(\tau_M) = e^{[(-\mathbf{R'}+\mathbf{K})\tau_m]}\mathbf{M'}(0) \qquad \textit{Eqn 5.2}$$

where **M′** equals:

$$\begin{bmatrix} \mathbf{M}_L^b \\ \mathbf{M}_L^f \\ \mathbf{M}_P^b \\ \mathbf{M}_P^f \end{bmatrix} \qquad \textit{Eqn 5.3}$$

where the superscripts b and f refer to the bound an free species, respectively, and the subscripts P and L refer to protein and ligand respectively. **R′** is an expanded relaxation matrix contains matrixes for the complexed ('b') and the uncomplexed ('f') states in the block-diagonal form:

$$\mathbf{R}' = \begin{bmatrix} \mathbf{R}_{(\tau_{cb}, r_{IS})} & 0 \\ 0 & \mathbf{R}_{(\tau_{cf}, r'_{IS})} \end{bmatrix} \qquad \textit{Eqn 5.4}$$

where τ_{cb} and τ_{cf} are the rotational correlation times describing the two states and r_{IS} and r'_{IS} are the corresponding internuclear distances. The exchange matrix **K** has the form

$$\mathbf{K} = \begin{bmatrix} -p_f k\mathbf{I} & p_b k\mathbf{I} \\ p_f k\mathbf{I} & -p_b k\mathbf{I} \end{bmatrix} \qquad \textit{Eqn 5.5}$$

where **I** is the identity matrix, and k is the reduced rate constant such that $1/\tau_b = k_{bf} = k_{off} = p_f k = (1\text{-}p_b)k$, and $1/\tau_f = p_b k1$. The forward rate constant is explicitly dependent on the free protein concentration according to $k_{fb} = k_{on}[P] = k_{on}([P_0] - [LP])$. The form of the kinetic matrix is determined by the fact that the exchange process connects only the corresponding nuclei of the free and bound ligand. Thus, the complete matrix governing the relaxation behaviour of the system are given by

$$-[\mathbf{R}'+\mathbf{K}] = \begin{bmatrix} \mathbf{R}_{(\tau_{cb}, r_{IS})} + p_f k\mathbf{I} & -p_b k\mathbf{I} \\ -p_f k\mathbf{I} & \mathbf{R}_{(\tau_{cf}, r'_{IS})} + p_b k\mathbf{I} \end{bmatrix} \qquad \textit{Eqn 5.6}$$

Intensities of the TRNOESY cross-peaks can be calculated according to equation 5.2 using the normalised initial matrix:

$$\mathbf{M}(0) = \begin{bmatrix} p_b\mathbf{I} & 0 \\ 0 & p_f\mathbf{I} \end{bmatrix} \qquad \textit{Eqn 5.7}$$

Thus, the theoretical transferred NOEs for the WGA-oligoG_{M3} can be computed using 'MDNOE2' by inputting the following variables:

[Protein]	0.644 mM (based on 4 sites per dimer)
[Sugar]	5.444 mM
K_D	1 mM (Kronis and Carver, 1985)
k	1000 s^{-1} (London *et al.* 1992)
$\tau_{c\ protein}$ (τ_{cb})	27 ns
$\tau_{c\ sugar}$ (τ_{cf})	0.2 ns (Berg *et al.* 1989)
τ_M	75 ms

where [Protein] and [Sugar] are the concentrations of WGA and oligo-G_{M3}, respectively, K_D is the equilibrium dissociation constant, k is the reduced rate constant, $\tau_{c\ protein}$, and $\tau_{c\ sugar}$ are the correlation times of WGA and the sugar respectively, and τ_M is the TRNOESY mixing time.

5.3 Results and Discussion

Solution Structure and Dynamics of [U-^{13}C] Oligo-G_{M3} in Free Solution

Complete proton and carbon chemical shifts for [U-^{13}C] oligo-G_{M3} at 300K were obtained from 2D HCCH-COSY (*Figure 5.2A*), 2D ^{13}C-^{13}C COSY (*Figure 5.2B*), and 2D ^{13}C-^{1}H HSQC (*Figure 5.3A*) experiments and assignments (See Table 5.1) were consistent with the partial proton (Siebert *et al.* 1992) and carbon assignments (Sabesan and Paulson, 1986) previously reported for this sugar. It is clear from Table 1 that there is significant spectral overlap in both the carbon and proton dimensions. Spectral overlap in the proton dimension is a particular problem as the interpretation of ^{1}H - ^{1}H NOESY and ROESY experiments requires the unambiguous identification of cross-peaks and the accurate measurement of cross-peak volumes for structural determination. To illustrate this point, Figure 5.4A shows a 1D slice through the Gal H1 4.56 ppm from the 2D ROESY spectrum of unlabelled oligo-G_{M3}.

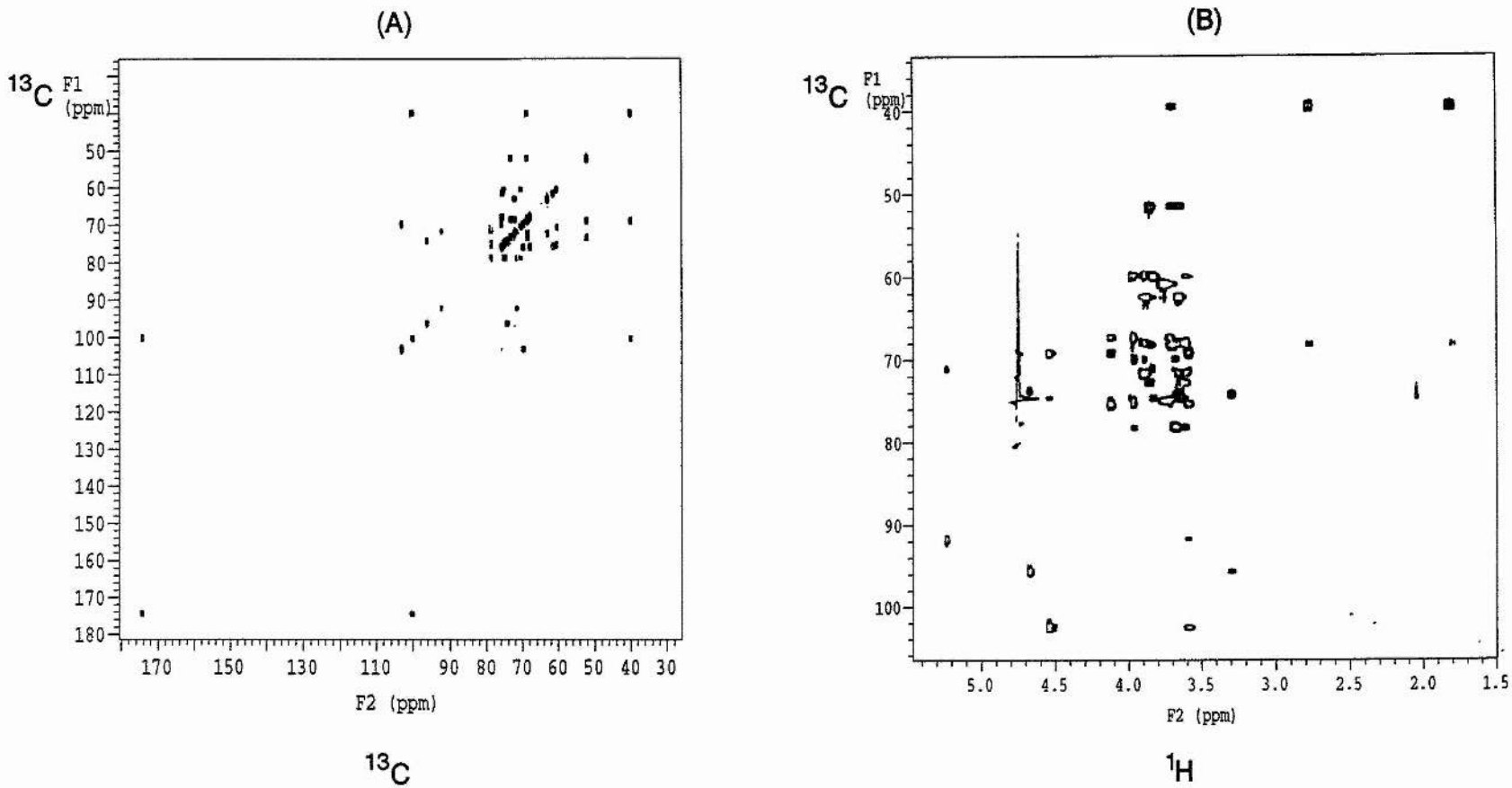


Figure 5.2: (A) 2D ^{13}C-^{13}C COSY and (B) HCCH-COSY spectrum of oligo-G_{M3} at 300K.

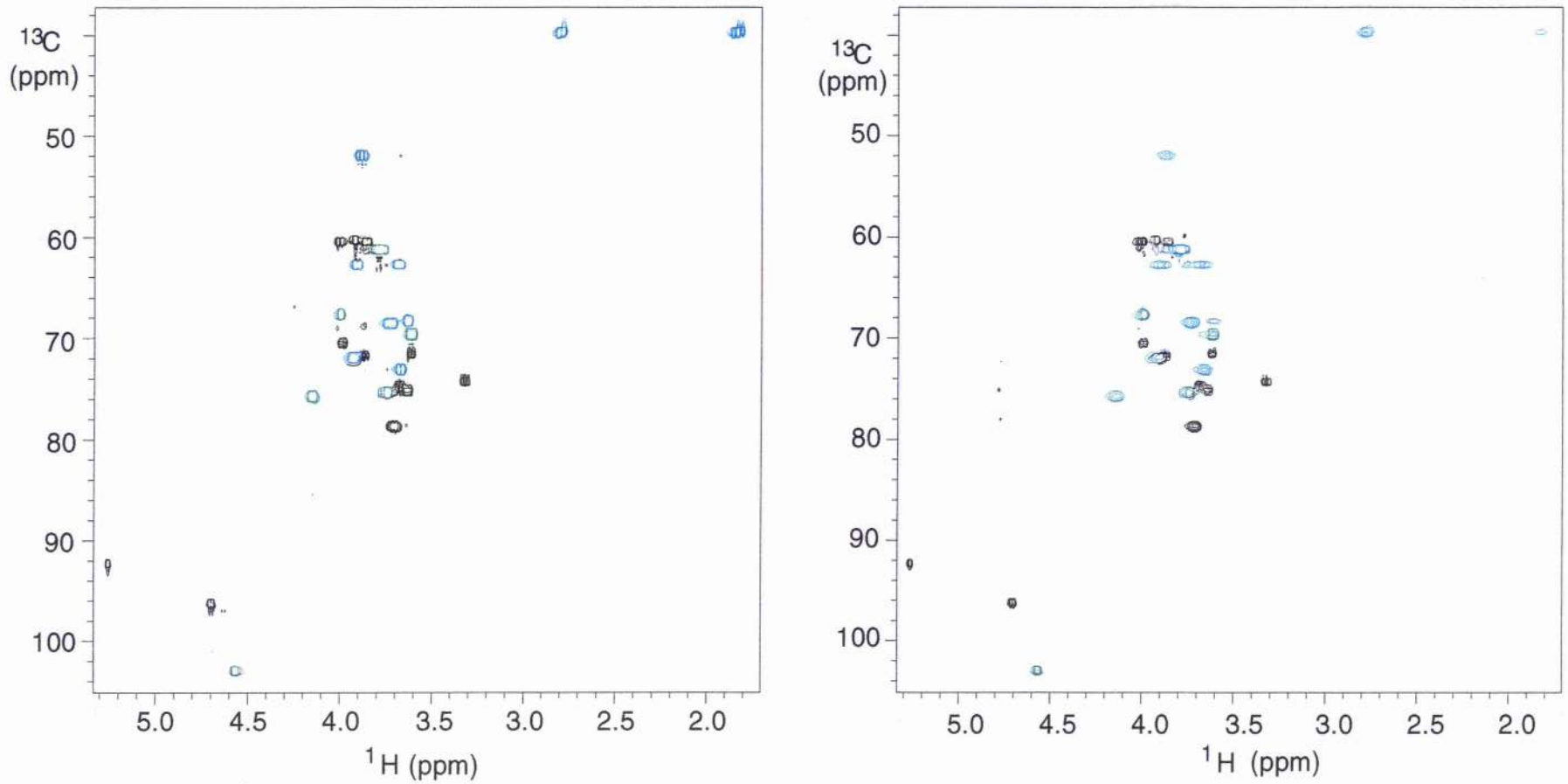


Figure 5.3: Heteronuclear ^{13}C-^{1}H correlation (HSQC) spectrum of (A): oligo-G_{M3} in free solution. (B): 5.44 mM oligo-G_{M3} in the presence of 0.161 mM WGA. Spectra were recorded at 300K. Peaks in Green: Gal, Cyan: Neu5Ac, and Black: α/β Glc.

Residue	Position	Proton(ppm)		Carbon(ppm)
		Free	Bound	
Neu5Ac	1	-	-	174.3
	2	-	-	100.3
	3_{ax}	1.84	1.83	40.1
	3_{eq}	2.79	2.78	-
	4	3.72	3.72	69.0
	5	3.88	3.86	52.3
	6	3.66	3.66	73.4
	7	3.63	3.60	68.7
	8	3.92	3.90	72.4
	9_{proS}	3.90	3.90	63.2
	9_{proR}	3.68	3.66	-
	CH_3	2.06	2.05	
Gal	1	4.56	4.56	103.2
	2	3.60	3.60	70.0
	3	4.14	4.12	76.2
	4	3.99	3.98	68.1
	5	3.74	3.74	75.8
	6_{proS}	3.78	3.78	61.6
	6_{proR}	3.78	3.78	-
βGlc	1	4.69	4.69	96.5
	2	3.31	3.31	74.5
	3	3.67	3.67	75.0
	4	3.69	3.69	79.0
	5	3.64	3.64	75.4
	6_{proS}	4.00	4.00	60.8
	6_{proR}	3.85	3.85	-
αGlc	1	5.25	5.25	92.6
	2	3.61	3.61	71.9
	3	3.88	3.88	71.9
	4	3.69	3.69	79.0
	5	3.98	3.98	70.7
	6_{proS}	3.91	3.91	60.6
	6_{proR}	3.91	3.91	-

Table 5.1: Chemical shifts for oligo-G_{M3}. "Free" solution shifts measured at 300K at pD 7.0. "Bound" chemical shifts measured for a solution containing 5.44 mM oligo-G_{M3}, 0.161 mM WGA in 0.1 M NaPi - 0.15 M NaCl at pD 6.4. All shifts were referenced indirectly to TSP.

A few cross-peaks may be unambiguously identified e.g. 4.14 ppm-Gal H3, 4.00/3.85 ppm-βGlc $H6_{proR}/H6_{proS}$, 3.78 ppm-Gal $H6_{R/S}$ but several cross-peaks which may describe the conformation about two of the glycosidic linkages could not be uniquely assigned e.g. 3.91 ppm could be a NOE from Gal H1 to Neu5Ac H8 (as reported in (Siebert *et al.* 1992)) and/or to both αGlc $H6_{R/S}$. The quantification of the ROE to Glc H4 is also difficult due to contamination with ROE intensities arising from Gal H5/H6s and possibly to βGlc H3/H5. The observed crosspeaks were separated out by their ^{13}C chemical shift in the 3D ROESY-HSQC experiment. An *F2F3* 2D slice through the Gal H1 resonance (*See Figure 5.4B*) reveals a well resolved ^{13}C-^{1}H correlation spectrum which allows the unambiguous assignment of all ROE connectivities to Gal H1. It is clear from this experiment that the proposed Gal H1 - Neu5Ac H8 (^{1}H 3.91 ppm, ^{13}C 72.42 ppm) (Siebert *et al.* 1992) is in fact an ROE cross-peak to αGal H6 (^{1}H 3.91 ppm, ^{13}C 60.62 ppm). Two additional ROEs are resolved in the 3D experiment. Figure 5.4B clearly shows ROE connectivities to α/β H3 and α/β H5. These ROEs have not been reported for oligo-G_{M3} in the literature. The primary reason these cross-peaks were overlooked in earlier studies is obvious with reference to Figure 5.4A. Figure 5.4A is a 1 D slice through the homonuclear 2D ROESY spectrum and it is clear that Glc H3/5 cross-peaks are completely swamped by the more intense intra-residue cross-peaks, in addition it would be impossible to separate out these cross-peaks using any combination of multidimensional homonuclear experiments.

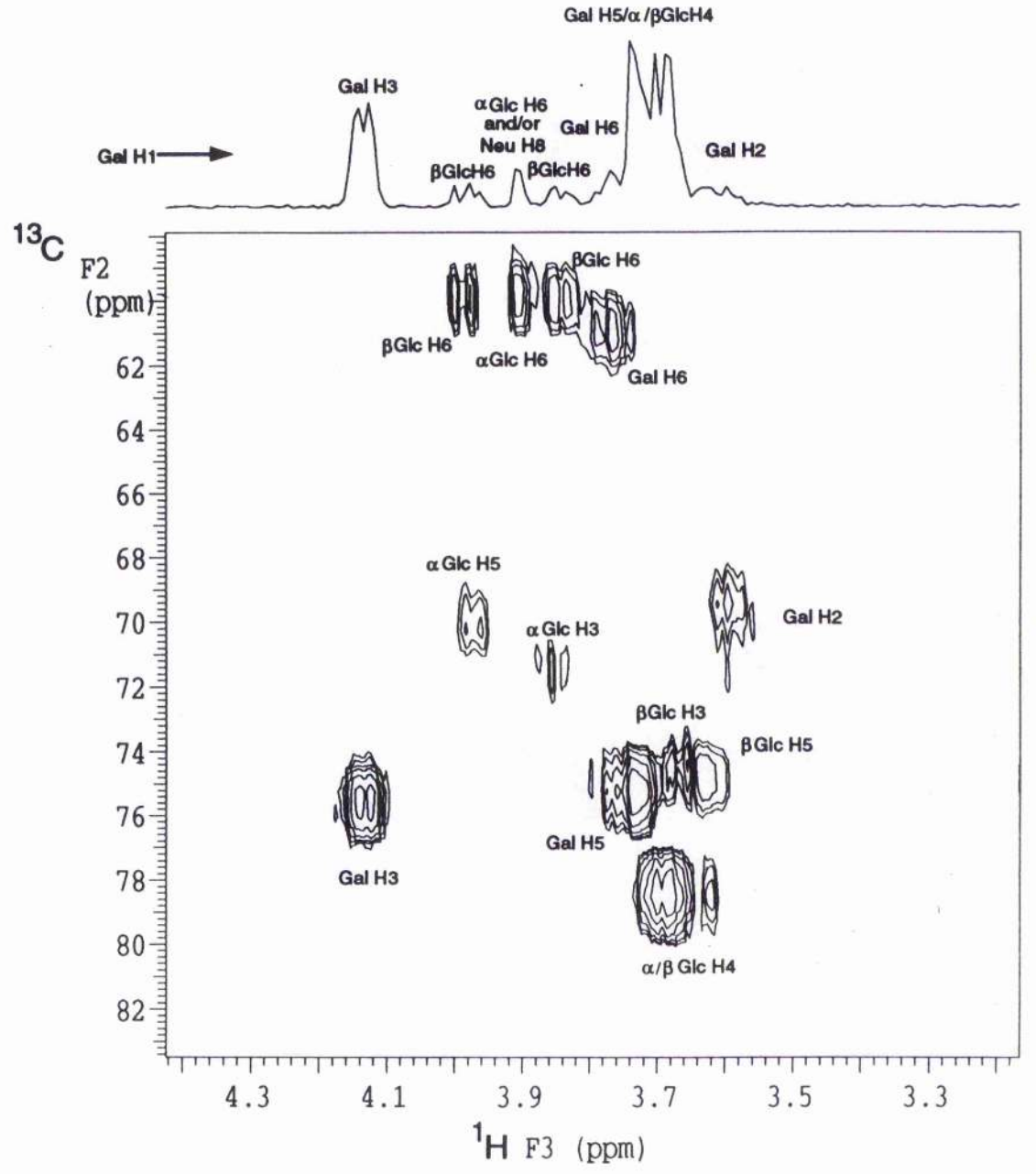


Figure 5.4(A): A 1D slice through Gal H1 (4.56 ppm) of a 2D ROESY spectrum. Initial ROE assignments based on the homonuclear experiment are labelled. (B): A *F2F3* (^{13}C, ^{1}H) slice through Gal H1 of a 3D ROESY-HSQC spectrum of [U-^{13}C] oligo-G_{M3}.

A *F1 F3* slice through 76.00 ppm (Gal C3/5 βGlc C3/5) (*See Figure 5.5*) highlights ROEs from Gal H3 to Neu5Ac $H3_{axial}$, and Neu5Ac H8. ROEs from βGlc H3 and H5 back to Gal H1 are also illustrated.

No other ROEs could be observed in the 3D experiment, and we could find no evidence for the Gal H4 to Neu5Ac H8 ROE as reported in (Siebert *et al.* 1992). All inter-residue ROEs were used as distance restraints in molecular modelling simulations (*See Table 5.2).*

Table 5.2: Experimental ROEs vs. Theoretical ROEs back-calculated from 5ns Restrained Dynamics Simulations.

Proton Pair	Relative ROE[b]	Theoretical ROE	r(Å[a])	Constraint
Neu $H3_{ax}$ - Gal H3	0.46	0.66	2.75	W
Neu H8 - Gal H3	0.16	0.25	3.28	W
Gal H1 - βGlc H4	1.87	1.73	2.18	S
Gal H1 - βGlc $H6_{proR}$	0.19	0.25	3.19	W
Gal H1 - βGlc $H6_{proS}$	0.20	0.30	3.16	W
Gal H1 - βGlc H3	0.18	0.19	3.22	W
Gal H1 - βGlc H5	0.20	0.27	3.16	W

[a]Approximate effective ^1H-^1H distance, based upon a simple ROE ratio calculation which neglects the influence of internal motions and anisotropic reorientation.

[b]Reference ROE Gal H1-H3, reference distance 2.43 Å.

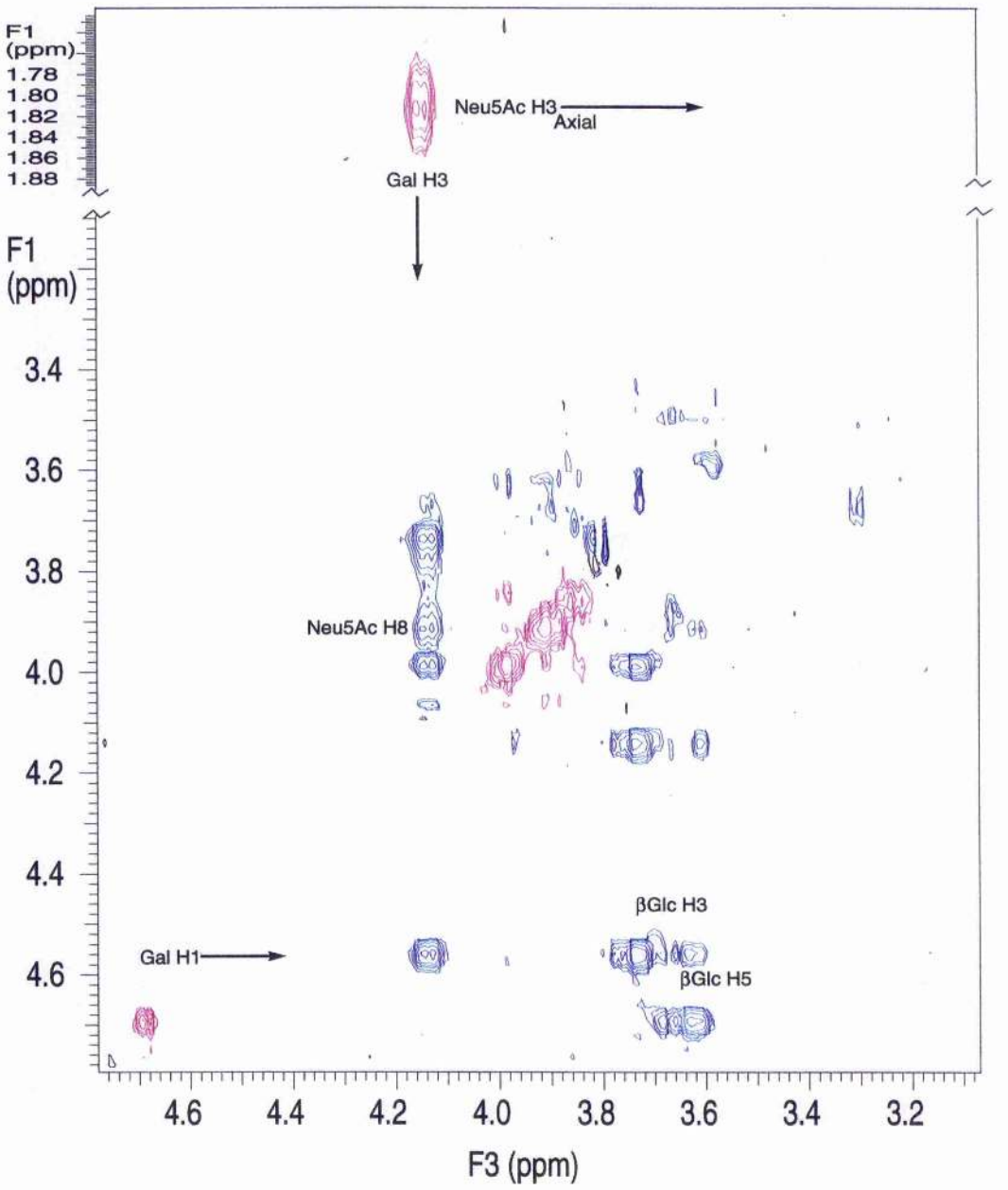


Figure 5.5: A *F1F3* (^{1}H-^{1}H) slice through the 3D ROESY-HSQC at Gal C3 (76.2 ppm).

Modelling results

The solution conformation and dynamics of intact G_{M3} and oligo-G_{M3} has previously been reported by (Siebert *et al.* 1992) using distance mapping procedures based on inter-residue ROE contacts which included ROE's between Neu5Ac H8 to Gal H1 and Gal H4. However, we could not find any evidence for these ROEs from the 3D ROESY-HSQC experiments (*vide supra*) performed on the carbon-13 enriched sample, therefore, a new investigation into the solution structure and dynamics of this carbohydrate was undertaken.

Table 5.2 lists the restraints used in restrained simulated annealing and MD simulations. The Neu5Acα2-3Gal linkage has been the subject of many NMR and modelling studies. Berg and co-workers (1989) first noted that the Neu5Acα2-3Gal moiety of α2,3 sialyl *N*-acetyl lactosamine may adopt 3 different conformations in free aqueous solution (*See Figure 5.6*): conformer 'A' ϕ,ψ -70°, 5°; conformer 'B' ϕ,ψ -165°, -20; and conformer 'C' -95°, -45°. Similar solution minima have been observed for the sialyl Lewis ^{x} moiety (Rutherford *et al.* 1994). Siebert and co-workers (1992) study of G_{M3} and oligo-G_{M3} predicted both conformers 'A' and 'B' plus an additional conformer '3' (ϕ, ψ 84°, 34°, respectively), whilst G_{M3} in magnetically orientated membrane appears to adopt conformer 'A' only (Aubin *et al.* 1993). Conformer 'B' is characterised by an ROE from Gal H3 to Neu5Ac $H3_{axial}$ and 'A' is defined by an ROE from Gal H3 to Neu5Ac H8. Both of these ROEs are observed in the 3D ROESY-HSQC, thus, oligo-G_{M3} must be in conformational exchange in free solution.

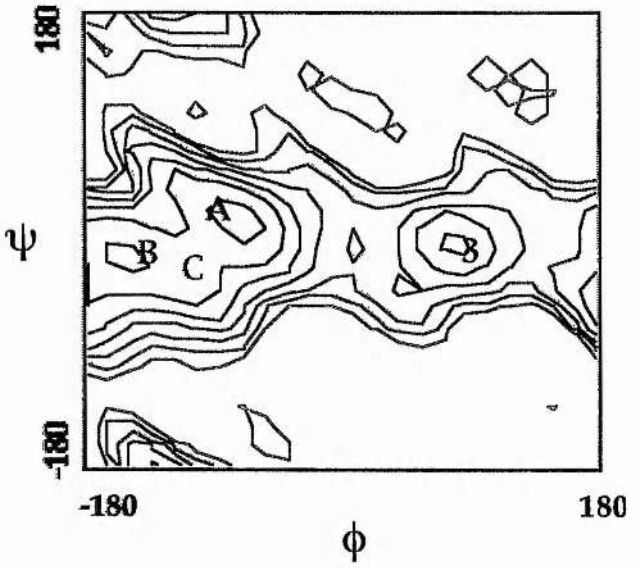


Figure 5.6: Potential energy surface for the Neu5Acα2-3Gal linkage showing the solution conformers 'A' and 'B' predicted by Berg (1989) and the additional solution minima predicted by Siebert (1990). Contours are in steps of 4.2 kJ mol^{-1}.

Preliminary NOESY experiments performed on oligo-G_{M3} in H_2O/15% acetone-d_6 at low temperature (data not shown) also generate conflicting NOE restraints. NOEs can be clearly seen between Gal OH2 and Neu5Ac $H3_{axial}$, and Gal H3 and Neu5Ac OH9. Ten computer models of oligo-G_{M3} were constructed, with pseudo-random geometries generated from an unrestrained high-temperature (750K) MD simulation (with all torsional terms scaled by a factor of seven to prevent unrealistic ring distortions). Each of these starting structures was optimised using the restrained simulated annealing protocol described previously (Rutherford 1993). The structures were annealed using ROE derived distance restraints listed in Table 5.2. Five of the ten starting structures converged to conformer 'B' (including the global minimum structure), three annealed to conformer 'A' and one to conformer 'C' (*Figure 5.7*, Table 5.3).

The remaining structure violated the Gal H1-Glc H4 restraint and was discarded. The lowest energy structure of each family was used as an input structure for a restrained MD simulation. Previous restrained dynamics simulations performed on sialyl Lewis ^{x} antigen highlighted the need for long MD simulations (up to 5 ns) as oscillations between conformers 'A' and 'B' are slow on the MD timescale (≤ 12 per ns). 5 ns of restrained MD simulations were performed on oligo-G_{M3}, each of the simulations gave similar results so only the global minimum is discussed further. Figure 5.7 shows the trajectory profiles for oligo-G_{M3} over the 5 ns simulation. Conformers 'A', 'B', and 'C' are sampled in solution (indicated by synchronous changes in ϕ and ψ values) and back calculated ROEs are in good agreement with experimental values (See Table 5.2).

Table 5.3: Minimised structures from restrained simulated annealing.

	Neu5Acα2-3Gal		Glycerol			Galβ1-4Glc	
Conformer	ϕ	ψ	ϕ_1	ϕ_2	ϕ_3	ϕ	ψ
B	-152.9	-25.0	-60.3	-178.3	-179.3	37.3	0.0
B	-148.9	-23.1	-62.1	-176.7	50.5	49.8	-1.9
B	-159.5	-16.8	-57.9	-175.5	-58.0	37.8	-0.4
B	-153.8	-23.7	-57.8	-175.5	51.4	38.5	-2.7
B	-155.0	-20.2	-59.7	-177.8	59.0	48.6	-0.7
A	-77.8	12.8	-57.9	-175.8	50.4	48.7	2.9
A	-81.4	15.5	-56.7	-175.0	-51.5	36.2	13.3
A	-81.3	15.8	-59.1	179.7	179.5	46.6	6.4
C	-119.9	-42.8	-56.5	-176.5	-57.2	-11.1	-28.2

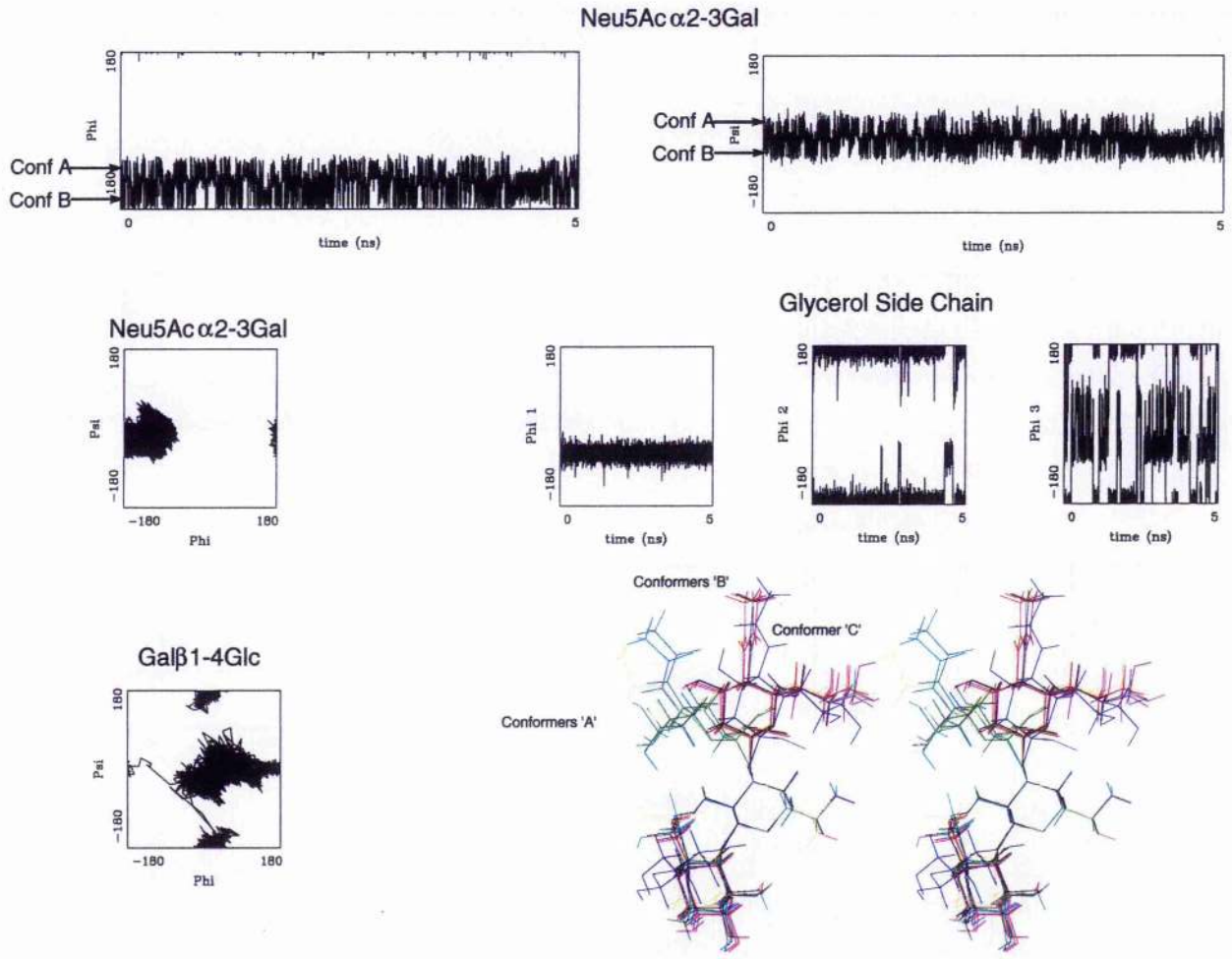


Figure 5.7: Phi vs. Psi trajectory plots for the 5 ns restrained dynamics simulations performed on oligo-G_{M3}. (Bottom right) Minimised structures from restrained simulated annealing protocol.

Conformation of the Galβ1-4Glc linkage

The 3D ROESY-HSQC experiment revealed two previously unreported ROE constraints across the Galβ1-4Glc linkage for oligo-G_{M3}, although similar NOEs have been observed for α2,3 sialyl *N*-acetyl lactosamine at low temperatures (Berg *et al* 1989). The ROE Gal H1 correlating α/β Glc H3/H5 is illustrated in Figure 5.5. Identical experiments performed on [U-^{13}C] Neu5Acα2-3Galβ1-4GlcNAC has yielded very similar results (Harris, 1997, *personal communication*). In contrast, 3D ROESY-HSQC experiments performed on [U-^{13}C] lactose or *N*-acetyl [U-^{13}C] lactosamine showed no dipolar interactions between Gal H1 and α/β Glc H3/H5 (Harris, 1997, *personal communication*). These data suggest that the presence of a Neu5Ac residue appears to subtly change the flexibility of the Galβ1-4Glc linkage. For an appreciable ROE to build up for between Gal H1 to α/βGlc H3/H5 this linkage must adopt the 'anti' conformer (ϕ, ψ 0°, -180°) for a period of time. It is clear from the trajectory plot for this linkage (*See Figure 5.7*) that the molecule does spend approximately 10% of its lifetime in this local minimum. Back-calculated ROEs for Gal H1 to βGlc H3/H5 (Table 5.2) are also in excellent aggreement with experiment indicating that the model of molecular motion predicted by the simulation is valid.

Spin coupling constants

Long range carbon-carbon and carbon-proton coupling constants are a potentially rich source of additional structural information. Long-range carbon-proton couplings were measured (See Table 5.4) from the constant

time experiment developed by Bax (1994) (data not shown). Also listed in Table 5.4 are back-calculated $^{3}J_{CH}$ values from the 5 ns restrained dynamics simulations calculated using an in-house written software 'MDPROCESS' (written by S. W. Homans) with a trans-glycosidic Karplus relationship derived by Tvaroska (1989). Values are in excellent agreement with those measured experimentally (experimental error ± 0.4 Hz).

Table 5.4 Experimental vs. Theoretical Inter-residue $^{3}J_{CH}$ values.

Linkage	Torsion	Experimental[a]	Theoretical
Neu5Acα2-3Gal	ψ	4.7	4.5
Galβ1-4Glc	ϕ	3.9	3.7
	ψ	5.1	4.9

[a] Experimental error ± 0.4 Hz

Long range J_{cc} values may be measured the 2D LRCC spectrum (Bax, A. *et al.* 1992), however, oligo-G_{M3} suffers from poor ^{13}C spectral dispersion. Figure 5.8A shows an extract from the 2D LRCC spectrum and highlights the problems with ^{13}C spectral overlap, $^{3}J_{cc}$ values from Gal C1 to βGlc C3/5 could not be measured because of poor digital resolution in *F1* and contamination with $^{2}J_{cc}$ cross-peak to Gal C3. The only conformationally relevant $^{3}J_{cc}$ value directly measurable from the 2D spectrum was between Neu5Ac C3 and Gal C3. In the 3D experiment a slice through the *F2 F3* dimension at 75.00 ppm (Gal C3/C5 βGlc C3/C5) results in a 2D ^{13}C-edited LRCC spectrum (*See Figure 5.8B*), to which $^{3}J_{cc}$ values can easily be

measured between βGlc C5 to Gal C1. No $^{3}J_{cc}$ cross-peak was observed between α/β Gal C3 to Gal C1 suggesting a $^{3}J_{cc}$ value of < ~1.0 Hz (A similar observation has been reported for $^{3}J_{C1C3}$ value in *N*-acetyl [U-^{13}C] lactosamine)(Harris, 1997, *personal communication*). A total of 4 inter-residue $^{3}J_{cc}$ values could be clearly measured (*See Table 5.5*) with an additional $^{3}J_{cc}$ between α/β Glc C4 to Gal C2 resolved in the 3D spectrum (*Shown in Figure 5.9*). Theoretical coupling constants were computed from the 5 ns dynamics simulation with a preliminary Karplus relationship for trans-glycosidic carbon-carbon couplings (Milton *et al* 1997). Table 5.5 lists the the back-calculated $^{3}J_{cc}$ values from the 5 ns dynamics simulations which are in general agreement with those measured experimentally (approximate experimental error ± 0.5 Hz).

The long-range spin-coupling constants are necessary for assessing computer models, as its information that is independent of ROE data. This is particularly important in the present case, as on the basis of ROEs and dynamics trajectories, we suggest that the 'anti' conformation (ϕ, ψ 0°, 180°) about the Galβ1-4Glc linkage must be significantly populated in solution. Dabrowski and colleagues (1995) claimed to have observed the 'anti' conformation for sugars dissolved in DMSO, however, this current study is the first (to our knowledge) example of the existence of the 'anti' conformation in aqueous solution. Thus, the excellent agreement between theoretical and experimental $^{3}J_{CC}$, and $^{3}J_{CH}$ values adds weight to the proposed existence of the 'anti' conformation.

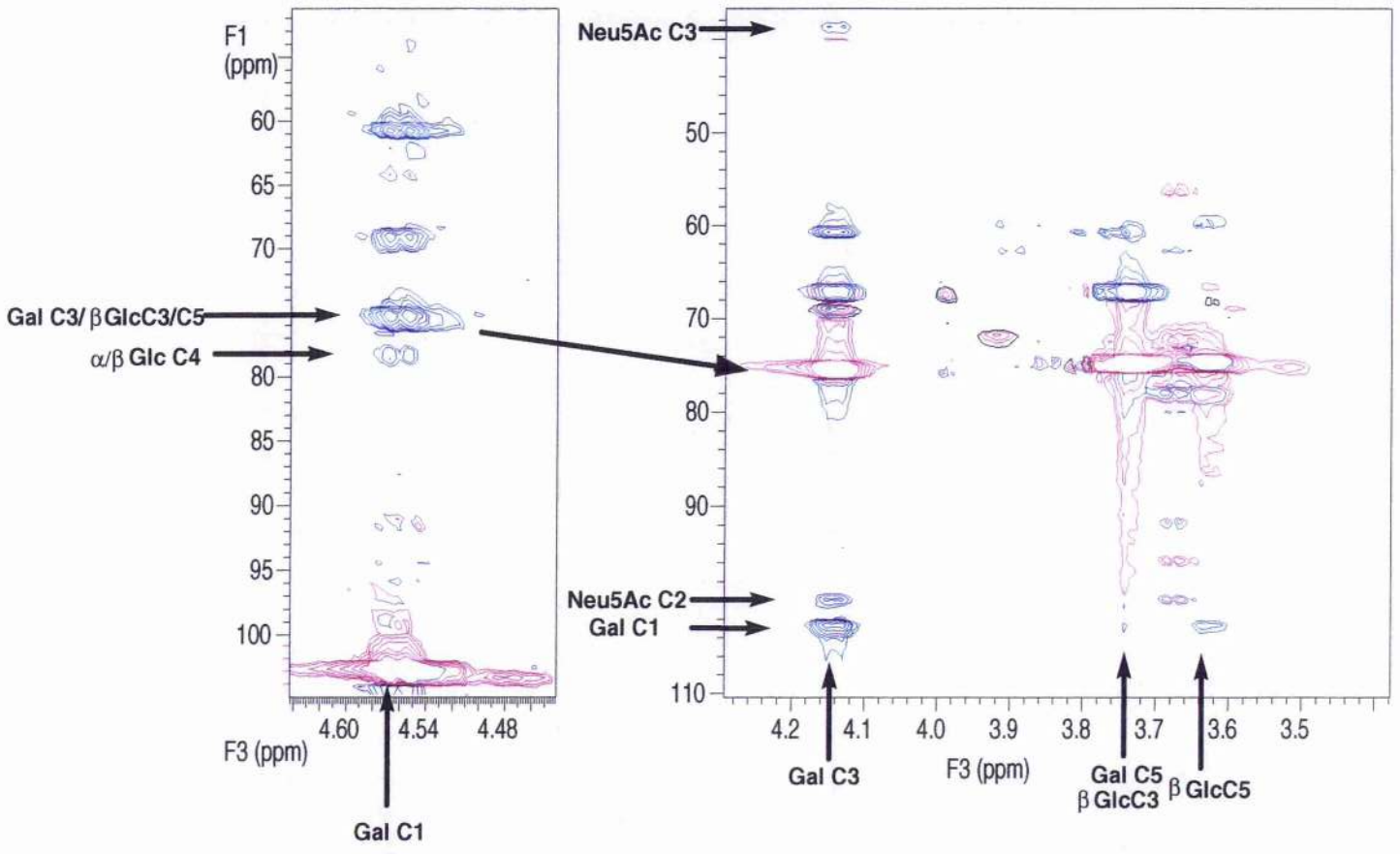


Figure 5.8(A): An extract from the 2D LRCC spectrum showing a cross-peak from Gal C1 to possibly Gal C5 and βGlc C5/3 (B): This cross-peak is expanded in the F1 dimension in the 3D experiment, the resolved cross-peaks from Gal C5 and βGlc C5 back to Gal C1 can now be clearly measured. Peaks in RED are reference peaks (V_{Ir}) and peaks in BLUE are long range cross-peaks (V_{ref}).

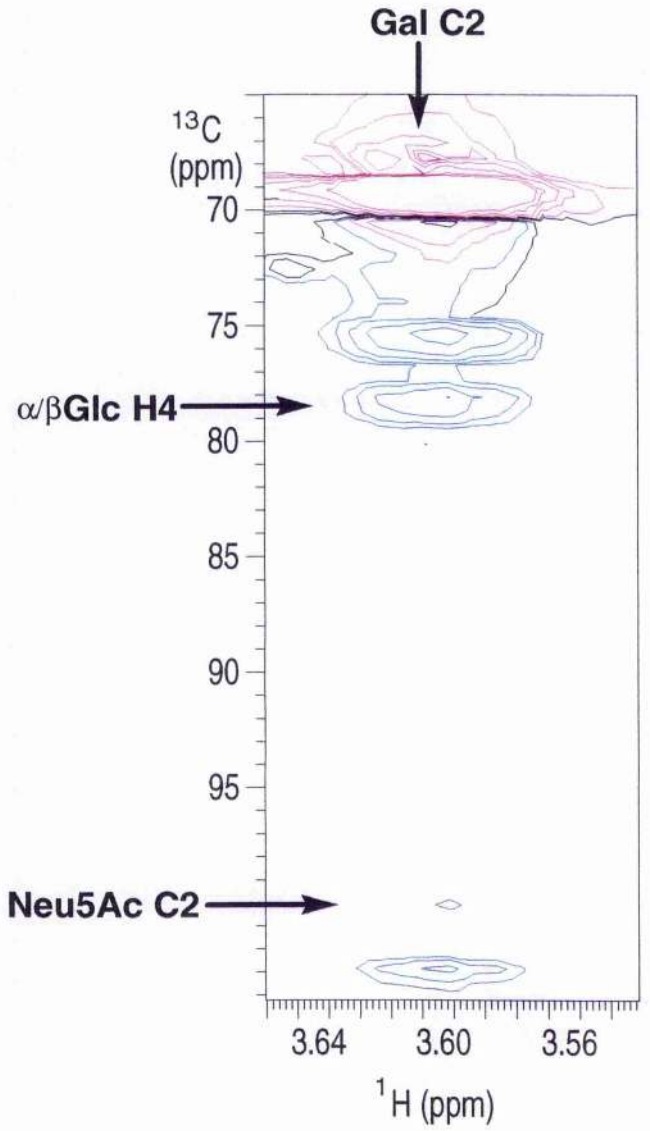
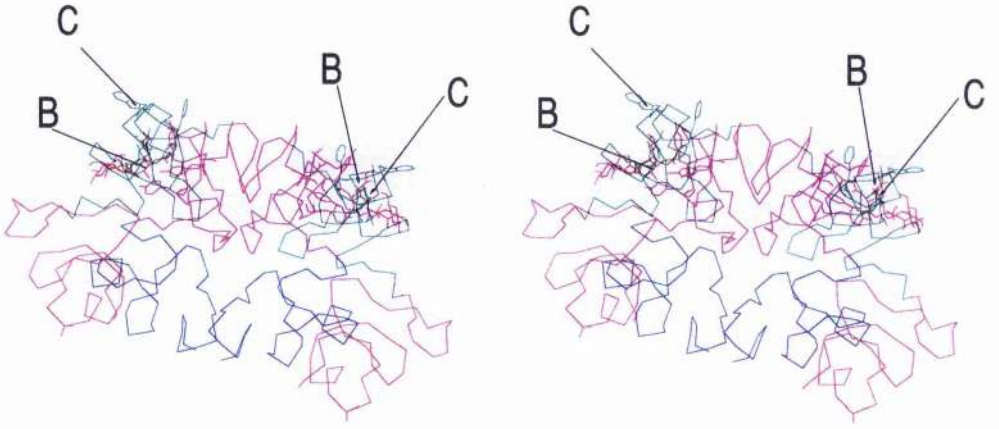


Table 5.5: Experimental vs. back calculated trans-glycosidic long range carbon-carbon coupling constants for labelled oligo-G_{M3} at 300 K. Experimental error +/- 0.5 Hz

$^{3}J_{cc}$	Experimental	Theoretical
Gal C2 - Neu5Ac C2	1.8 Hz	1.8 Hz
Gal C3 - Neu5Ac C3	1.9 Hz	1.6 Hz
Gal C4 - Neu5Ac C2	< 1.4 Hz	1.1 Hz
Gal C1 - βGlc C5	1.9 Hz	1.6 Hz
Gal C2 - α/β Glc C4	2.6 Hz	2.5 Hz
Gal C1 - α or βGlc C3	< 1.4 Hz	1.1 Hz

Solution Conformation of Oligo-G_{M3} in Complex with WGA

Wheat germ agglutinin is a highly stable disulfide-rich lectin produced by polyhaploid wheat (*Triticum vulgaris*) which specifically interacts with GlcNAc and Neu5Ac terminating oligosaccharides. The WGA monomer consists of four similar 43-amino acid residue domains (A, B, C, D) arranged in tandem, and dimerises in a "head to tail" fashion, forming an extensive monomer/monomer interface (*See Figure 5.10*). The number of oligosaccharide binding sites proposed for the dimer has ranged from 2-12. Up to 8 sites per dimer have been identified by X-ray crystal crystallography (four unique sites related by the dimers two-fold symmetry (Wright, 1992), however, NMR (Kronis and Carver, 1982; Kronis and Carver, 1985) and microcalorimetric studies (Bains, *et al.* 1992) of WGA in complex with small ligands performed in free solution suggest that only 4 of the 8 binding sites are functional in the non-cooperative binding of mono-, di-, and trisaccharides (*vide infra*). Binding sites are located at contact points between dimers and are composed of two binding regions: a shallow pocket characterised by three quasi-conserved aromatic amino acids, and one conserved serine on one domain; and a nonconserved region that consists of one or two polar residues, on the contacting domain of the second monomer. In the nomenclature proposed by Wright (1996), each unique site is referred to by the domain that contributes the aromatic pocket, thus in the X-ray crystal structure the eight binding sites are described as $A1_{(D2)}$ or $A2_{(D1)}$, $B1_{(C2)}$ or $B2_{(C1)}$, $C1_{(B2)}$ or $C2_{(B1)}$, $D1_{(A2)}$ or $D2_{(A1)}$ respectively. Theoretical calculations performed by Wright and Kellogg (1996) suggest that the four "active" sites detected in free solution studies are $B1_{(C2)}$ and $B2_{(C1)}$, and $C1_{(B2)}$ and $C2_{(B1)}$.



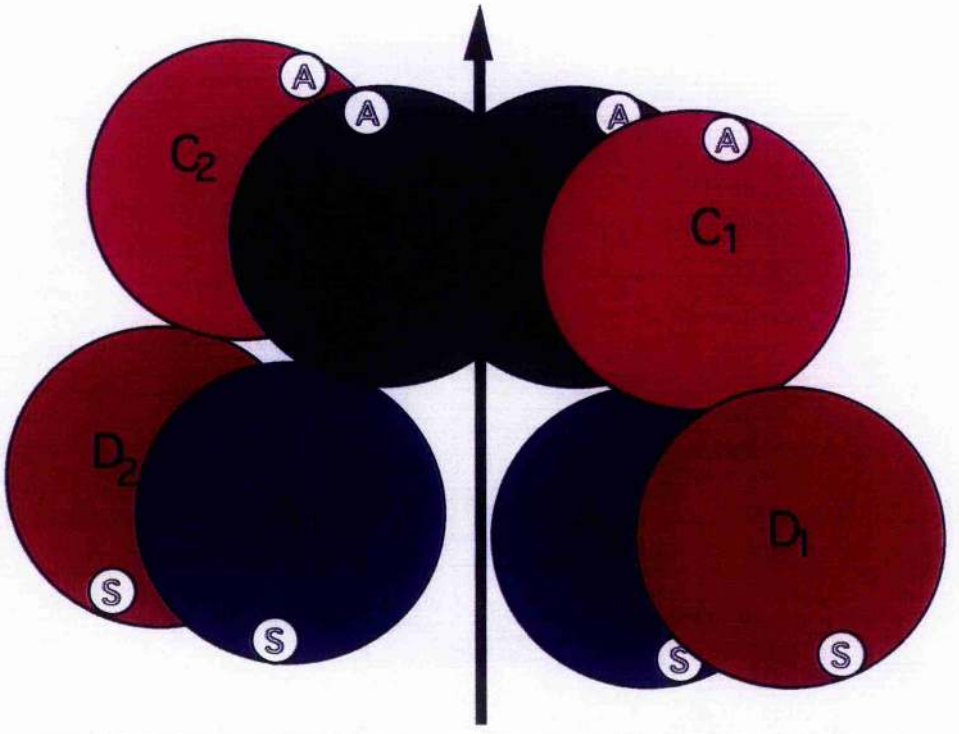


Figure 5.10(A): X-ray crystal structure of WGA in complex with oligo-G_{M3} (Wright, C. S. 1990; Wright, C. S. and Kellog, G. E. 1996) highlighting the 'active' B and C sites in solution. (B): Schematic representation of the WGA dimer showing all eight binding sites identified by X-ray crystallography (Wright, C. S. 1992). (A) 'active' site (S) secondary site.

Modelling studies predict that the D sites do not bind Neu5Ac due to unfavourable charge interactions, and the A sites are thought to be nonfunctional in solution due to the small number of H-bonds that A sites can make with Neu5Ac.

Commercial WGA samples have been reported to contain between 3 (Rice & Etzler (1975)) to 9 isolectins (Matsumoto *et al.* 1987). Isolectins 1 (WGA1) and 2 (WGA2) are the dominant species in protein preparations and make up ~35% and ~49% of the purified weight of WGA, respectively. Isolectins are thought to be a result of multiple gene copies present in polyhaploid wheat and all isolectins share high sequence homology (95-97%). WGA1 and 2 differ in only 5 amino-acid positions, one being located in the ligand binding site $B1_{(C2)}$ (or $B2_{(C1)}$) (WGA1 Tyr66 to His66 in WGA2) and effects the binding affinity for sialyltrisaccharides with WGA1 binding sialyloligosaccharides more tightly than WGA2. The difference in K_D's between the 2 isolectins (~0.8 mM and ~3 mM, for WGA1 and 2, respectively) is thought to arise from the more favourable interactions that occur when residue 66 is a tyrosine (Wright, 1984; Wright, 1990). The tyrosyl side-chain provides a larger surface area for van der Waals' stacking against the Neu5Ac pyranose ring than His66 and there is one less H-bond possible in the WGA2 $B1_{(C2)}$ site. No other amino-acid substitutions in the proposed binding sites have been observed for other isolectins. Oligo-G_{M3} has only been observed bound to the two B sites in WGA1 and WGA2 (Wright, 1984; Wright, 1990). The α2,3 arm of of the T5 antigen (Neu5Acα2-3Galβ1-3GalNAcα6-(Thr)-2Neu5Ac) complexed in the $C2_{(B1)}$ binding pocket has also been elucidated by crystallography (Wright, 1992). Theoretical TRNOEs for the

oligo-G_{M3} - WGA system were calculated (based on the X-ray coordinates of the B and C binding sites) to determine if the solid state models gave a valid description of the bound conformation of oligo-G_{M3} in solution.

The selective line-broadening and shifting of several resonances in the proton spectrum of oligo-G_{M3} in the presence of WGA has been noted by Kronis and co-workers (Kronis and Carver, 1982; Kronis and Carver, 1985) and is illustrated in Figure 5.3B. Line broadening effects due to increases in solution viscosity or non-specific binding have been ruled out by experiments performed by Kronis (1982) and no resonance perturbations could be seen in the proton spectrum of oligo-G_{M3} in the presence of bovine serum albumin (BSA) (data not shown). Intramolecular and intermolecular transferred NOEs have been reported by Umemoto and co-workers (1988) for 1-*O*-methyl GlcNAc in the presence of WGA: selective irradiation of sugar protons in free solution gave small positive NOEs, when the same experiment was performed on a sample containing WGA there was a change in sign of NOEs (small and negative) which is indicative of NOEs building up in the slowly tumbling bound state. An identical observation was made when selective 1D transient NOEs experiments were performed on unlabelled oligo-G_{M3} in the presence of WGA (data not shown). This confirmed that useful distance information on the bound conformation of oligo-G_{M3} may be extracted from 2D TRNOESY experiments.

Optimising experimental parameters to measure TRNOEs.

To minimise partial nulling of negative TRNOEs from positive NOEs arising from the free sugar (which is in great excess), selective 1D transient NOE experiments were performed at various temperatures

(between 18° and 30 °C) on a standard sample of unlabelled oligo-G_{M3}. Strong intra-residue NOEs reached a null at 27 °C.

In order to determine the best experimental parameters for measuring TRNOEs, 0.5 mg aliquots of WGA was added to a stock solution of unlabelled oligo-G_{M3} (2.5 mM) at 27 °C. The magnitude of TRNOEs was monitored by recording 2D TRNOESY experiments after each addition of WGA. TRNOEs reached maximum intensity at a ligand:protein concentration of 33:1 (based on protein molecular weight of 36 kDa). A mixing time of 75 ms was found to be optimal from such studies. The effects of spin diffusion are a major source of erroneous distance information in TRNOE studies (Arepalli *et al.* 1995), however, spin diffusion is manifested as cross-peaks with positive intensities in transferred ROESY (TRROESY) spectra. 3D TRROESY-HSQC spectra (data not shown) performed on [U-^{13}C] oligo-G_{M3} - WGA sample showed no evidence of cross-peaks due to spin diffusion.

In common with ROESY experiments on oligo-G_{M3} in free solution, analysis of TRNOESY experiments is hampered due to proton spectral overlap, interpretation of TRNOESY experiments is further complicated by line-broadening, chemical shift perturbations due to ring current shifts (Kronis and Carver, 1985), and contamination from protein signals. These problems were alleviated by employing a carbon-13 enriched ligand. Proton and carbon chemical shifts for the exchanging ligand were measured from a ^{13}C-^{1}H HSQC spectrum and are listed in table 5.1. It is clear from table 5.1 that Neu5Ac $H3_{equ}$/$H3_{axial}$, H5/H7/H8/$H9_{pro\text{-}R}$ and Gal H3/4 are all shifted in the presence of WGA. These protons also

experience significant line broadening. Line broadening is primarily due to chemical exchange where ligand protons which form intimate contacts with the protein experience a different magnetic environment in the bound state (See Chapter 1, Equation 1.15)(Swift and Connick, 1962). In contrast line shapes for the α/β glucose resonances experience no significant line broadening or chemical shift perturbations at the protein concentrations used for TRNOESY experiments (*see Figures 5.3A & B*) indicating that this residue must have a similar mobility to the free solution state and make little contact with the protein. These observations have important implications when interpreting theoretical TRNOEs. It is obvious from the *F1F3* projection of the 3D TRNOESY experiment (*Figure 5.11*) that there is no apparent NOE cross-peak correlating Gal H3 (4.14 ppm) to Neu5Ac $H3_{axial}$ (1.82 ppm) in the bound state. A (TR)NOE from Gal H3 to Neu5Ac $H3_{axial}$ is characteristic of conformer 'B' (*vide supra*), it is therefore likely that only a subset of the conformational hyperspace accessed in solution is sampled in the bound state. Its not immediately clear from the *F1 F3* projection (equivalent to the information available from a 2D homonuclear experiment) that there is a TRNOE between Gal H3 and Neu5Ac H8, which is characteristic of conformer 'A'. A *F2F3* 2D plane through Neu5Ac H8 (3.91 ppm)(*Figure 5.12*) clearly shows a small TRNOE providing evidence that conformer A is selected on binding as predicted from X-ray crystallography. However, the intensity of the TRNOE is difficult to quantify in the *F1F3* dimension due to the marked line broadening that these two resonances experience.

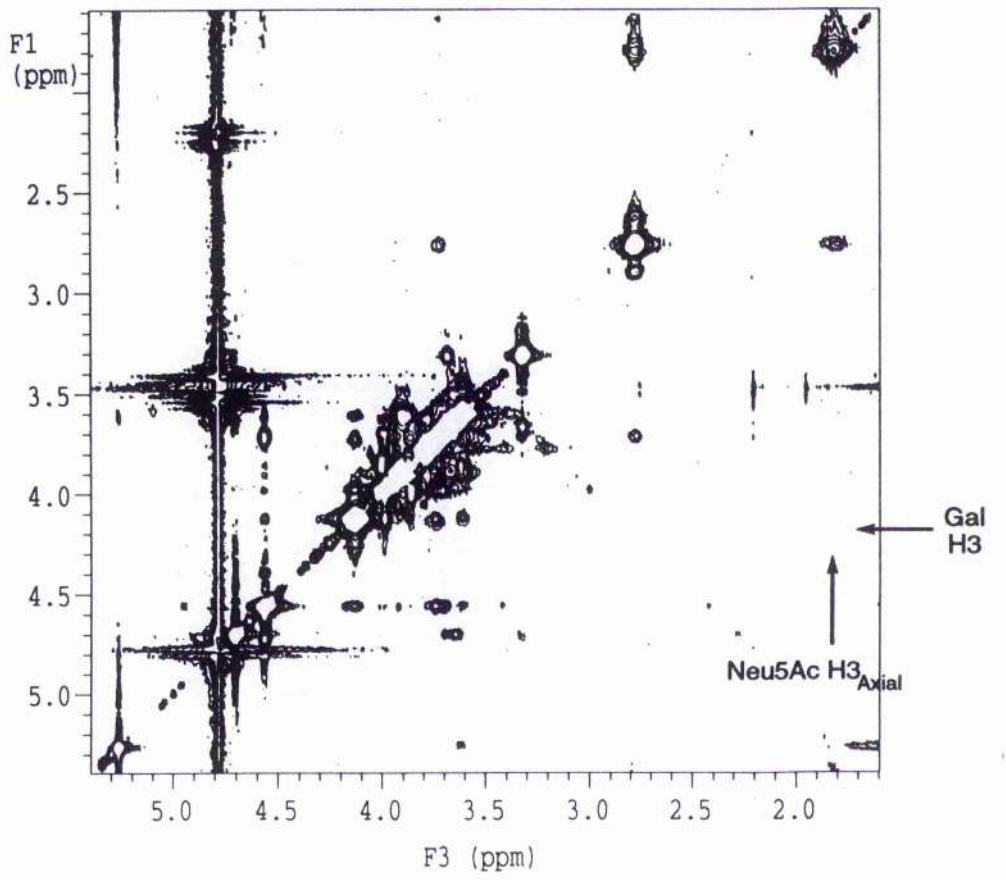


Figure 5.11: A *F1F3* projection of the 3D TRNOESY-HSQC Spectrum. Note there is no Gal H3 - Neu5Ac $H3_{axial}$ cross-peak.

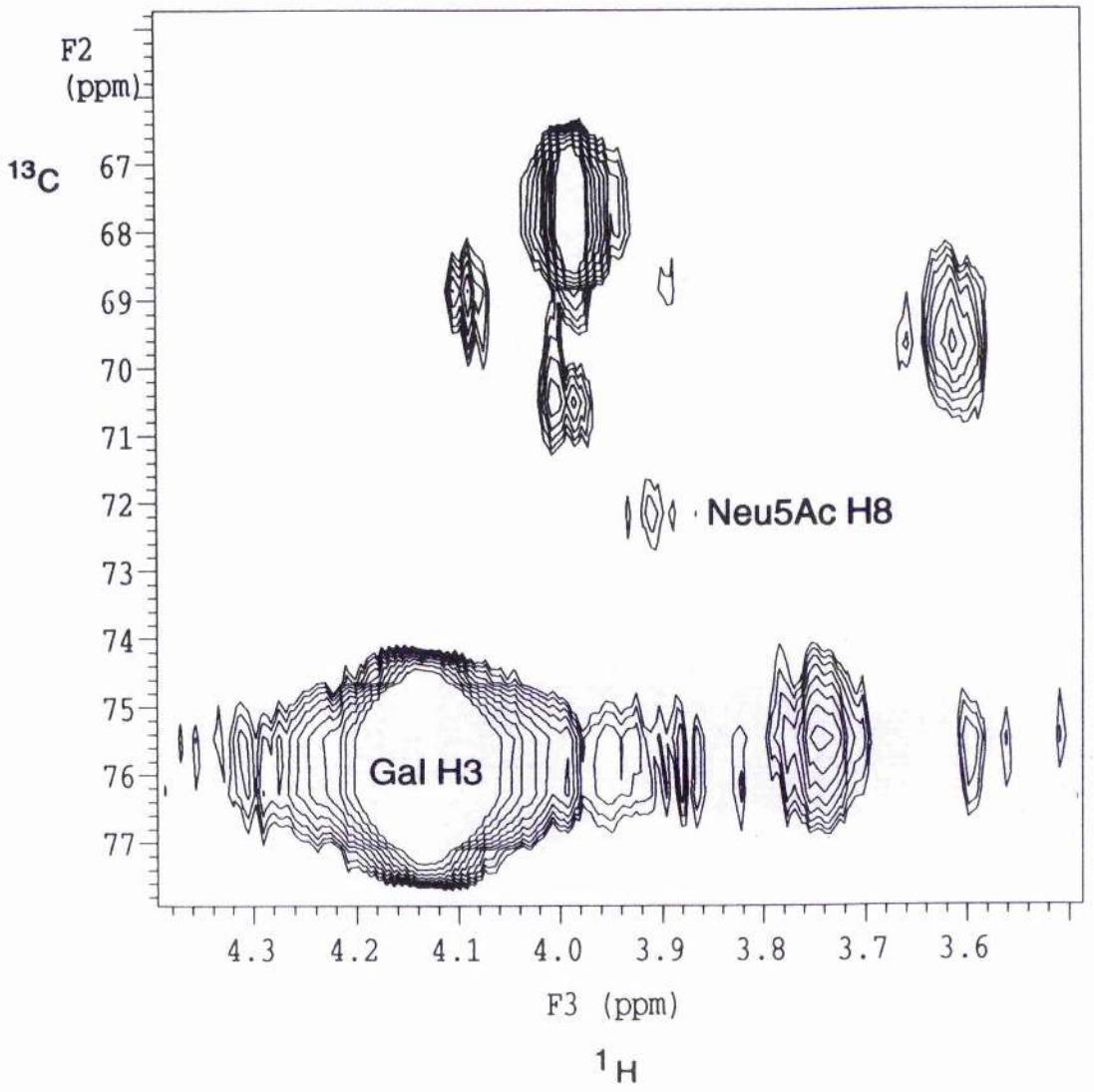


Figure 5.12: A *F2F3* slice through the 3D TRNOESY-HSQC at Gal H3 (4.12 ppm).

Additional NMR restraints are listed in Table 5.6.

Table 5.6: Observed inter-residue TRNOEs and results from simulated annealing.

TRNOE restraints used in restrained simulated annealing

Proton pair	Restraint	
Gal H1 - Glc H4	M	1.8-3.3Å
Gal H1 - Glc $H6_{pro\,S}$	W	1.8-5.0Å
Gal H1 - Glc $H6_{proR}$	W	
Gal H3 - Neu5Ac H8	W	

Conformer	ϕ	ψ	ϕ_1	ϕ_2	ϕ_3	ϕ	ψ
A	-74	13	-54	179	-55	44	3
A	-72	9	-56	174	-40	36	0
A	-75	18	-57	180	50	45	1
A	-84	18	-54	-169	53	8	-13
A	-82	13	-57	-171	52	42	2
A	-74	13	-53	-174	53	7	-18
A	-69	13	-54	-178	-54	45	4
A	-80	17	-56	-171	53	43	3
A	-71	12	-59	-176	-49	48	4
A	-65	9	-60	-179	-50	40	1

A series of distance restraints were generated on the basis of these relative TRNOEs and were used in simulated dynamical annealing routines to determine if TRNOE based distance restraints alone could correctly predict the bound state conformation. 10 pseudo random starting geometries were subjected to restrained dynamical simulated annealing with the dielectric constant set to 4.00 to mimic the protein binding environment. All ten structures annealed to conformer 'A' about the Neu5Acα2-3Gal linkage and multiple conformers about the Galβ1-4Glc linkage. Conformer 'A' is the observed bound conformation in the X-ray crystal structure (listed in Table 5.7), although the predicted conformation

about the Galβ1-4Glc linkage is very different from those modelled in the X-ray structure. The electron density for the glucose moiety is very disordered which is characteristic of a highly mobile group in the crystal lattice, and may have been modelled incorrectly in the original study (Wright, 1990). Annealed torsional angles for the glycerol side chain of Neu5Ac also lie within the range of with those predicted by X-ray crystallography. These data suggest that it is possible to correctly predict the bound state conformation, on the basis of TRNOEs, of a linear sialyltrisaccharide without prior knowledge of the binding site structure.

Table 5.7: Structures used in TRNOE simulations.

Complexed structures of oligo-G_{M3} determined by X-ray crystallography								
Neu5Acα2-3Gal			Glycerol			Galβ1-4Glc		
Conformer	ϕ	ψ	ϕ_1	ϕ_2	ϕ_3	ϕ	ψ	Site
Model I	-69	-2	-90	-167	-35	98	-20	B1
Model II	-67	-12	-79	174	-35	92	-27	B2
Model III	-65	9	-61	162	-120	92	7	B1
Model IV	-72	1	-66	163	-88	52	-18	B2
'Computed' structures								
Model V	-69-	2	-90	-167	-35	40	0	B1
Model VI	-46	6	-57	-171	-79	24	-20	C2

Full relaxation matrix calculations

To test the validity of the X-ray crystal structures, as models for the bound conformation in solution, full relaxation matrix calculations were performed on models of the complex (listed in Table 5.7). One of the models used in these calculations is illustrated in figures 5.13. Full matrix relaxation calculations performed on this exchanging system must take

into account kinetic data such as dissociation constants, concentrations of ligand and complex, and relaxation properties of the complex and free ligand. These factors have been included in a full matrix relaxation algorithm for exchanging systems developed by London and co-workers (London *et al.* 1992). London's analysis was based on static ligand and complex architectures. As oligo-G_{M3} is clearly in conformational exchange between conformers 'A' and 'B' in solution, full relaxation matrix calculations based on a single fixed conformation for the free ligand would be inadequate. An extension of London's (1992) analysis to include TR(R)NOEs building up in free solution (based on the restrained 5 ns dynamics trajectory used to model theoretical ROEs) are included in an in-house written piece of software 'MDNOE2'.

There is only one weak TRNOE across the Neu5Acα2-3Gal glycosidic linkage and theoretical TRNOEs should reflect this if the X-ray crystal structures represent the dominant bound conformer in solution. Table 5.8 represents experimental vs. theoretical TRNOEs for the 6 models used in this study. TRNOEs between Neu H8 and Gal H3 match well with experiment for the structures derived for X-ray crystallography.

Table 5.8: Results from full-relaxation matrix calculations.

TRNOE	Experimental Relative TRNOE	Theoretical TRNOEs I	II	III	IV	V	VI
Neu5AcH8 - GalH3	0.16	0.11	0.15	0.19	0.32	0.11	0.18
Gal H1 - βGlc H4	1.76	0.27	0.17	0.17	0.24	1.65	1.67
Gal H1 - βGlc H6(R)	0.06	0.23	0.02	0.05	0.09	0.23	0.07
Gal H1 - βGlc H6(S)	0.08	0.07	0.15	0.05	0.05	0.07	0.50

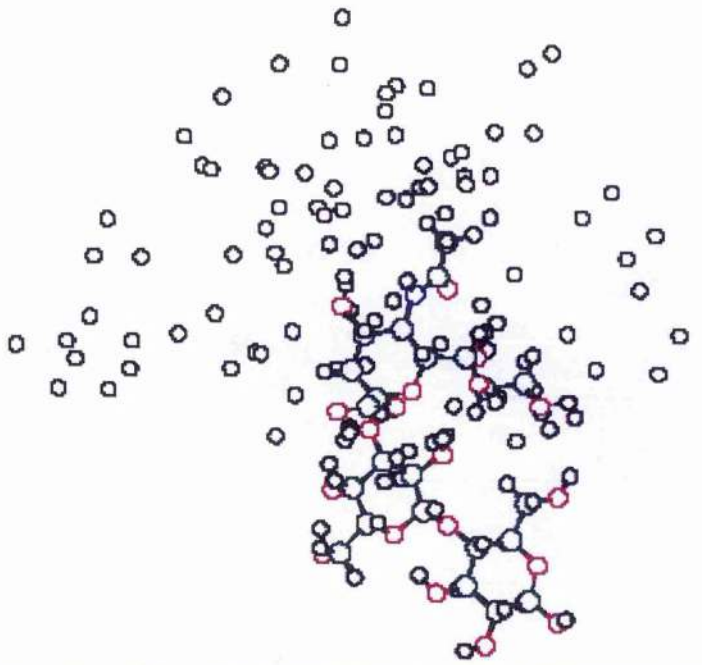


Figure 5.13: An example of the structures used to represent the WGA - oligo-G_{M3} complex in full relaxation matrix calculations. Isolated spheres represent protein non-exchangeable protons.

TRNOEs based on the X-ray crystal structures (models **I** to **IV**) clearly do not match observed TRNOEs for the Galβ1-4Glc linkage. This is not a surprising observation as the electron density for this moiety is very poor in the X-ray crystal structure (Wright, C. S. 1990). However, when the ϕ, ψ torsion angle for this linkage in model **I** (to give model **V**) was manually manipulated to ϕ, ψ 40°, 0°, respectively (this is the conformation for this linkage predicted by restrained simulated annealing conformation, and used as starting conformations in models **V** and **VI**) the theoretical relative TRNOEs approach those measured by NMR. In all the models, the two theoretical TRNOEs from Gal H1 to βGlc H6 pro-R and H6 pro-S could not be realised simultaneously. Similar results were also obtained for the $C1_{(B2)}$ site (model **VI**). It is clear from this study that ignoring the internal dynamics of a ligand when bound to the protein when

calculating theoretical TRNOEs may result in a poor agreement with experimental observations. In general the observations made by NMR match those predicted by X-ray crystallography, i.e. conformer 'A' is selected upon binding, with the Neu5Ac residue making extensive contacts in an aromatic rich region of the protein, with a highly mobile glucose moiety making little or no interactions with the protein.

5.4 Conclusions

To date only a few multidimensional heteronuclear NMR investigations of highly enriched oligosaccharide moieties has been undertaken. The oligo-G_{M3}-WGA system was chosen as a model system to test if heteronuclear NMR experiments had any advantages over their homonuclear counterparts, before the synthesis of more complex but biologically important *N*-acetyl [U-^{13}C] neuraminic acid containing saccharides is undertaken. It is clear that multidimensional heteronuclear NMR investigations of the solution behaviour of the oligosaccharide moiety of the ganglioside G_{M3} has distinct advantages over purely homonuclear (^{1}H) methods. Previous proton NMR investigations of this sugar incorrectly assigned some ROE connectivities and completely missed others. Complete unambiguous assignments of all (TR)NOESY/ROESY cross-peaks was possible and two additional ROEs were measured across the Galβ1-4Glc linkage. Additional structural information on the relative orientation about the glycosidic linkages were made from long range carbon-carbon and carbon-proton coupling constants. Restrained dynamics simulations performed on the sugar clearly show that this carbohydrate is exchanging between two dominant

solution conformations, this is in common with previous studies performed on sugars which terminate with the Neu5Acα2-3Gal moiety. $^3J_{CC}$ and $^3J_{CH}$ values back-calculated from dynamics simulations independently confirm that our computer models are consistent with experimental data.

Modelling studies appear to suggest that the 'anti' (ϕ,ψ 0°, 180°) conformer about the Galβ1-4Glc linkage is significantly populated in solution. Five theoretical ROEs and four spin-coupling constants were back-calculated for this linkage and were in excellent agreement with experimentally measured values. Thus, there is significant evidence for the adoption of this conformation in free solution. However, additional information from inter-residue NOE contacts between exchangeable and C-linked protons, may provide conclusive evidence for the adoption of the 'anti' conformer. Hardware limitations have so-far prohibited such studies.

The analysis of the WGA - oligo-G_{M3} complex was also aided with the application of multidimensional heteronuclear NMR spectroscopy. Theoretical full relaxation matrix calculations based on the X-ray crystal structures in exchange with conformers 'A' & 'B' in solution were performed. These data are in agreement with the experimental observations made from 3D ^{13}C-^1H TRNOESY-HSQC and and ^{13}C-^1H HSQC data, i.e. conformer 'A' is selected by both B and C sites upon binding to WGA, with the glucose residue making little contact with the protein, thus the Galβ1-4Glc linkage.

Chapter 6

A preliminary heteronuclear NMR study of N-Acetyl [U-^{13}C] neuraminic acid and [U-^{13}C] galactose labelled N-linked glycans of the Fc fragment of immunoglobulin G in situ.

Abstract

A preliminary heteronuclear NMR investigation of the dynamics of a *N*-linked complex glycan of an intact glycoprotein is reported. In an extension to the work published in (Gilhespy-Muskett *et al.* 1994) we have re-[U-^{13}C] galactosylated and [U^{13}-C] sialylated the glycan chains of a mutant Fc fragment (FcREA) of a immunoglobulin G. Two-dimensional heteronuclear ^{13}C-^{1}H HSQC experiments performed on the labelled glycans suggest that in common with galactose, the terminal sialic acid residue on the α1-3 arm experiences a different magnetic environment and mobility to the α1-6 arm. These data suggest that the lectin-like site on the protein surface proposed for galactose may extend to sialic acid.

6.1 Introduction

Immunoglobulin G (IgG) is a serum glycoprotein and is produced as part of the secondary humoral immune response. IgG is composed of two types of polypeptide chains: two heavy (H) and two light (L). The F_{ab} region is responsible for antigen binding, with the binding sites located in the Fv domain. The F_{ab} region is linked to the F_c domain by a flexible 'hinge' region which aids multivalent antigen binding. The F_c fragment consists of two constant H chains and this region is required for the activation the complement cascade and is recognised by Fcγ receptors on the surface of leucocytes.

IgG samples purified from human sera contain an asparagine linked oligosaccharide located in at Asn297 of the two heavy chains (Edelman*et al.* 1969) in the Fc region (*see Figure 6.1A*). *N*-linked glycans can also be found in the F_{ab} region of the H and L chains. Several studies have shown that the presence of the Asn297 glycan chain is essential for maintaining the functional structure of the C_H2 domain as removal of the glycans results in the loss of binding to Fcγ receptors; compliment binding and activation are compromised; and the molecule is more susceptible to proteolytic degradation.

Analysis of the glycosylation profiles of IgG samples obtained from healthy individuals reveals that the sugar chains are of the complex biantennary complex-type (shown in *Figure 6.1.B*), however, the composition of the glycans is very heterogeneous with most of the changes produced by the presence or absence of the two galactoses or the

'bisecting' *N*-acetyl glucosamine. Further structural heterogeneity arises from the differential sialylation of the terminal galactoses. Despite this seemingly bewildering range of carbohydrate chains (up to thirty-six different chains), the molar ratio of each oligosaccharide is constant among healthy individuals.

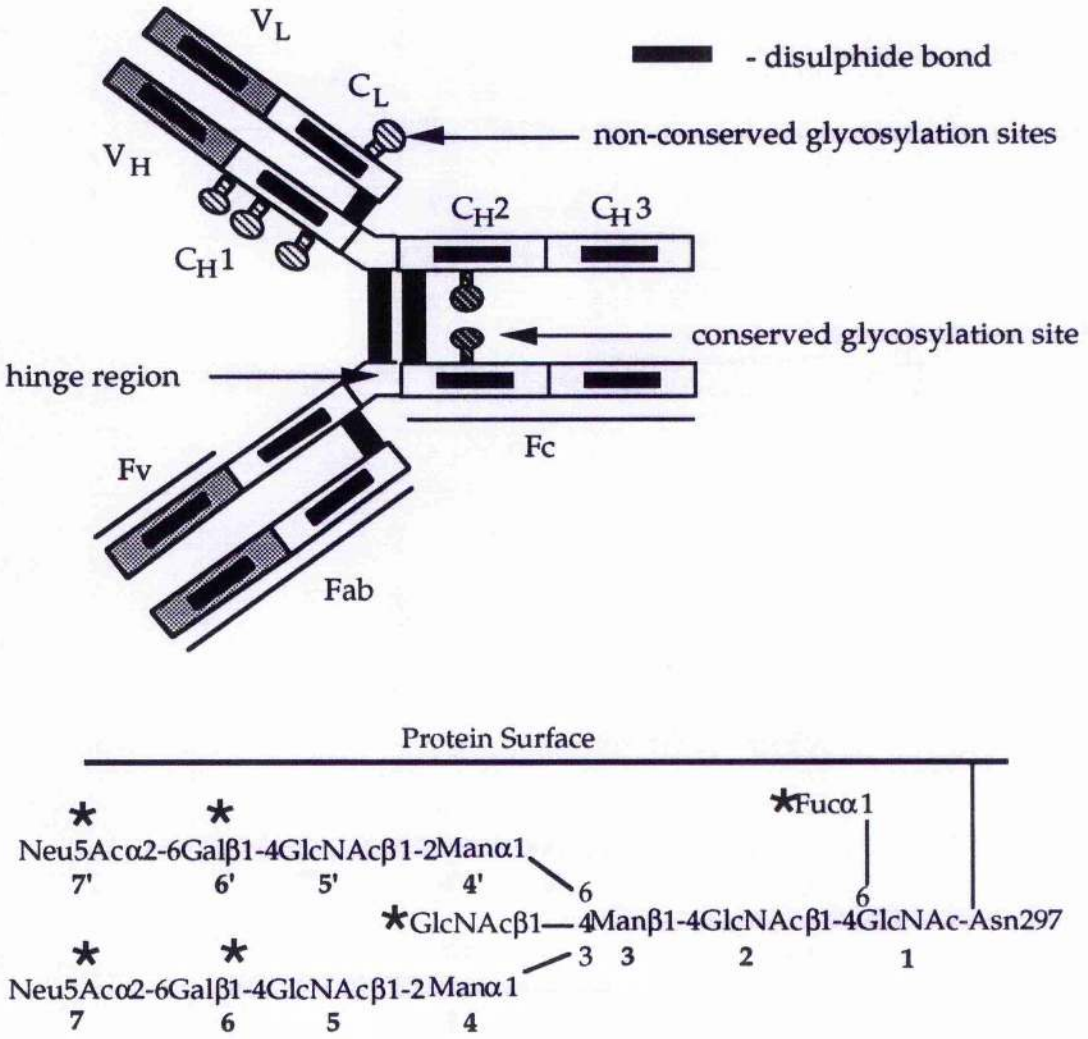


Figure 6.1: (Above) schematic representation of IgG showing the domain structure, pattern of disulfide bonds and glycosylation sites. (Below) structure of the glycan chain attached to Asn 297 of the C_H2 domain. The presence or absence of *residues represent the main sources of structural heterogeneity. Sugar chains typically terminate with galactose, although, up to 20% may be monosialylated and ~5% may be siaylated in healthy individuals (Parekh, R. B. 1985).

However, it soon became obvious that IgG extracted from patients with certain auto-immune diseases such as rheumatoid arthritis (RA), and Crohn's disease had altered glycosylation patterns. Patients with RA, for example, contained the same set of biantennary oligosaccharides found in normal individuals but a higher percentage of chains remained agalactosylated. These 'G0' glycoforms can be used both as an indicator of disease severity and in the early diagnosis of RA. The removal of the terminal galactoses is thought to expose a lectin like patch on the protein surface previously covered by galactose (Rademacher *et al.* 1985). These vacant sites could bind galactose residues from other IgG molecules, resulting in the formation of auto-aggregates typical of RA. Alternatively, the sites may reveal antigenic determinants, creating a direct auto-immune response. It is clear that the removal of the galactoses must effect the dynamics of the oligosaccharide chains and their interactions with the Fc fragment.

The aim of the current investigation is two-fold: to test the applicability of published methods for the incorporation of *N*-acetyl [U-^{13}C] neuraminic acid into an intact glycoprotein, and to investigate the influence of the terminating Neu5Ac residues on the structure and dynamics of the carbohydrate chain and its possible interactions with its native protein. It was also the hoped that we could study the conformation of the glycan chain when the Fc fragment was bound to an RA factor (a monoclonal antibody raised against a human Fc fragment).

6.2 Materials and Methods

Regalactosylation of FCREA

Lyophilised and salt free powder of FcREA (a mutant protein devoid of any terminal galactose residues) (16 mg) was dissolved into 50 mM Tris-HCl pH 7.4 (2.5 mL) containing 7 mM NaN_3, 5 mM $MgCl_2$, 20 mM $MnCl_2$, 10 mM ATP, 10 mM [U-^{13}C] D-galactose, galactokinase (5 U) (E.C. 2.7.1.6), and galactose-1-phosphate uridyltransferase (5 U) (E. C. 2.7.7.12), and galactosyltransferase (8U) (lactose synthase, E.C. 2.4.1.22). The reaction mixture was incubated at 37 °C overnight. [U-^{13}C] galactosylated FcREA was purified by protein G affinity chromatography.

Protein G Purification of Immunoglobulins

A protein G column (1 mL, 33 mg mL^{-1} binding capacity)(Sigma) was prepared by washing the column with 10 column volumes of 50 mM Tris-HCl pH 7.4, 3M KCNS pH 8.0, and once again with 50 mM Tris-HCl pH 7.4. The FcREA samples were dialysed into Tris-HCl pH 7.4 to a final buffer concentration of 50 mM.

FcREA was loaded on to the protein G column. The absorbance of the eluate was monitored at 280 nm (A_{280}). The column was washed with 50 mM Tris buffer until the A_{280} returned to zero (approx. 20 column volumes). FcREA was eluted of the column with 3M KCNS pH 8.0. 1 mL fractions were collected and samples containing protein were pooled (5 mL) and dialysed into phosphate buffer saline (PBS) to give a final buffer concentration of 10 mM (10 mM KPi, 30 mM NaCl, 28 mM KCl) pH 7.4. The sample was freeze-dried and dissolved in 1 mL D_2O and transferred

into a 5 mm NMR tube (528 pp, Wilmand).

Synthesis of α2-6 Sialylated IgG and [U-^{13}C] Galactosylated FcREA (Scheme 6)

[U-^{13}C] galactosylated FcREA, lyophilised salt-free powder, (6 mg), [U-^{13}C] Neu5Ac (~15 mg) and PEP (12 mg) was dissolved in HEPES buffer pH 7.5 (355 μL) containing 30 mM $MgCl_2$, 0.7 mM $MnCl_2$, 13 mM KCl, 14 mM CMP, 1.4 mM ATP, and 0.1% mecraproethanol. The pH of the solution was readjusted by step-wise addition of 2 M NaOH. Pyruvate kinase (PK, 8.5 U, E.C. 2.7.4.4), inorganic pyrophosphatase (PPase, 1 U, E.C. 3.6.1.1), nucleoside kinase (NMK, (0.5 U) E.C. 2.7.4.4), CMP-sialic acid synthase ((60 mU), E.C. 2.5.7.43), and sialyltransferase ((20 mU), E.C. 2.4.99.1). BSA (1 mg mL^{-1}) was added to protect FcREA against proteases. The production of CMP-Neu5Ac (R*f* 3.1) was easily monitored by TLC (1 M NH_4OAc/propan-2-ol 1:2.4). The reaction was left to incubate at room temperature for 3 days. To confirm the synthesis of α2-6 linked sialic acid, 2D ^{13}C-^{1}H HSQC experiments were performed on the crude reaction mixtures (*See Figure 6.3*).

NMR Spectroscopy

NMR experiments were performed at 37 °C or 65 °C. 2D ^{13}C-^{1}H HSQC experiments were recorded with proton and carbon sweep widths of 2 kHz of 12 kHz, respectively. 128 scans were acquired per t_1 increment. Spectra consisted of 128 complex points in t_1 and 1 K complex points in t_2.

6.3 Results and Discussion

Several NMR and X-ray crystallographic studies have been performed on intact glycoproteins to examine the role that the protein and glycan perform to maintain each others conformation. To date, most studies have indicated that the glycan chain extends out into solution and as a consequence makes little contact with its protein partner (Lommerse *et al.* 1995; Weller *et al* 1994, 1996; Wyss *et al* 1995). This is consistent with the proposed function of many glycoprotein glycans which is to act as a 'masking' agent preventing contact or tight binding with proteases or possible receptors, respectively. In contrast, X-ray crystallographic studies of Fc fragments of human (Deisenhofer, 1981) and murine (Harris *et al.* 1997) origin reveal that the sugar makes extensive contacts with the C_H2 domain. These observations are consistent with biochemical data which shows the essential role that the glycan chain plays in maintaining the active conformation of the C_H2 domain, which is required to bind Fcγ receptors and complement binding (Lund, 1990; Adler *et al.* 1995; Boyd *et al.* 1995), and the hypothesis that removal of the terminal galactose residues (G0 glycoforms) results in the exposure of a galactose binding site which may be responsible for the auto-aggregates associated with RA (Rademacher *et al.* 1985). Proton NMR relaxation studies performed on IgG glycoforms (Dwek *et al.* 1995) terminating with galactose were consistent with one arm of the glycan being 'bound' i.e. having the same relaxation properties as the protein, with the other arm extending out into the inter-domain space and moving more rapidly. When the terminal galactose residue 6′ was removed the relaxation properties of the glycan chain were indicative of a mobile group i.e. the sugar no-longer

made any intimate contacts with the protein and was free to move in the space between the Fc heavy chains. Heteronuclear NMR spectroscopy performed on a [U-^{13}C]-galactosylated Fc fragment (and illustrated in Figure 6.2) clearly shows a splitting and differential line broadening of the galactose cross-peaks. Homans and co-workers (Gilhespy-Muskett *et al.* 1994) suggested that this observation may confirm the hypothesis that α1-6 terminating galactose **6′** is 'bound' to a site at the C_H2 domain. Unlike other classes of immunoglobulins, only approximately 20% and 5% of human IgG isolated from the serum of healthy individuals may be mono- and di-sialylated, respectively (Parekh, 1985). The effect sialylation may have on the structure of the glycan chain and on the conformation of the C_H2 domain remains unclear. There are conflicting reports in the literature on the effects of sialylation. Adler and colleagues (1995) reported that asialylated IgG bound less effectively to FcγR receptors on the surface of polymorphonuclear leucocytes, whilst Boyd and co-workers (1995) reported no effects on IgG function after the removal of the terminal Neu5Ac residues. It is known, however, that masking the terminal galactose residues extends the serum half life of the glycoprotein (Newkirk *et al.* 1996). The steric effect of the Neu5Ac residue coupled with its negative charge may prevent the 'binding' of **6′** to the C_H2 domain, and it may repel the observed interactions between the α1-3 terminal galactose (**6**) residues. Following the successful synthesis of *N*-acetyl [U-^{13}C] neuraminic acid (discussed in chapter 4) it was now possible to incorporate labelled Neu5Ac and use heteronuclear NMR to probe its effects on the structure and mobility of the glycan chain.

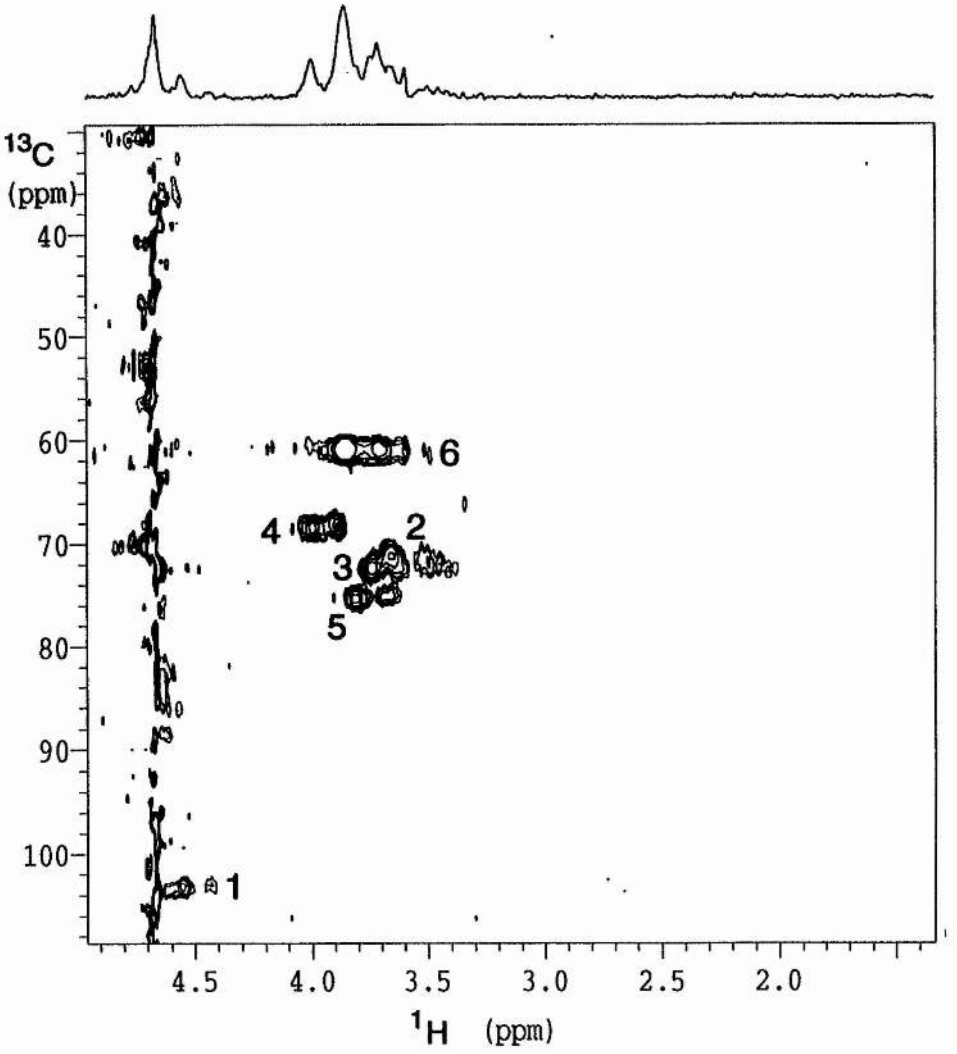
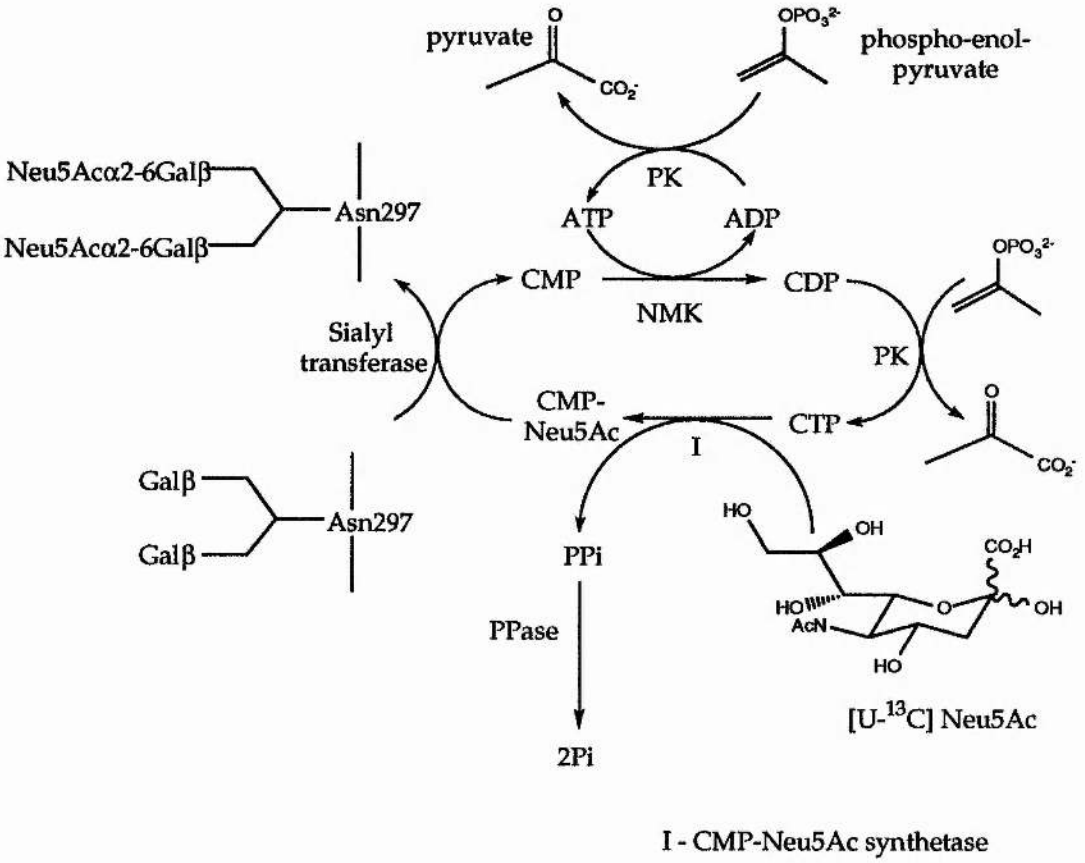


Figure 6.2 ^{13}C-^{1}H HSQC spectrum of [U-^{13}C] Gal labelled FcREA at 37 °C

Uniformly carbon-13 enriched galactose was incorporated into the FcREA glycans using *in-situ* generation of UDP-[U-^{13}C] galactose and galactosyltransferase as described previously (Gilhespy-Muskett *et al.* 1994). Galβ1-4GlcNAcβ1-R moieties may be α2,6 sialylated in high yield using commercially available α2,6 sialyltransferase and CMP-Neu5Ac, however, CMP-[U-^{13}C] Neu5Ac is not commercially available but may be synthesised from CTP, Neu5Ac, and CMP-sialic acid synthase. The equilibrium for this reaction is approx. 50 % and to push the reaction to completion, a typical strategy would be couple this reaction with the glycosyltransferase as we have with the synthesis of Galβ1-4GlcR oligosaccharides. Unfortunately, the CMP-sialic acid synthase reaction releases CDP which inhibits the sialyltransferase enzyme and results on poor yields of sialylated sugars. An alternative approach has been published by Wong and co-workers (1995) (*Scheme 6*). In this reaction scheme concentrations of nucleoside phosphates are kept to a minimum by *in situ* regeneration of nucleotide sugars Starting from Neu5Ac, LacNAc, phopspho enol pyruvate (PEP), and catalytic amounts of adenosine 5′-triphosphate (ATP) and cytosine 5′-monophosphate (CMP), with *in situ* regeneration of CMP-Neu5Ac, CTP, ATP, and ADP. Scheme 6 illustrates the reaction pathways, CMP was converted to CDP catalysed by nucleoside-monophosphate kinase (NMK) in the presence of ATP which is regenerated by pyruvate kinase (PK) in the presence of PEP. The CDP was further converted to CTP with PEP catalysed by PK. The CTP, then, reacted with Neu5Ac catalysed by CMP-Neu5Ac synthase to produce CMP-Neu5Ac. The by-product, inorganic pyrophosphate (PPi) was decomposed to inorganic phosphate by inorganic pyrophosphatase

(PPase). Sialylation of the *N*-linked glycans was accomplished by CMP-Neu5Ac and α2,6 sialyltransferase. The released CMP was again converted to CDP, to CTP, and to CMP-Neu5Ac. The recovery of FcREA form this reaction was approximately 70%.



I - CMP-Neu5Ac synthetase

Scheme 6: The multienzymatic synthesis of [U-^{13}C] sialylated FcREA. Abbreviations are listed in the text.

Figure 6.2B shows the C3/H3 cross-peaks for Neu5Ac. It is clear from this spectrum that as well as Neu5Ac, CMP-Neu5Ac and a α2,6 sialylated species (confirmed by comparison with published proton chemical shifts for these species (Schauer, 1982) and these species can unambiguously be identified by both their proton and carbon chemical shifts.

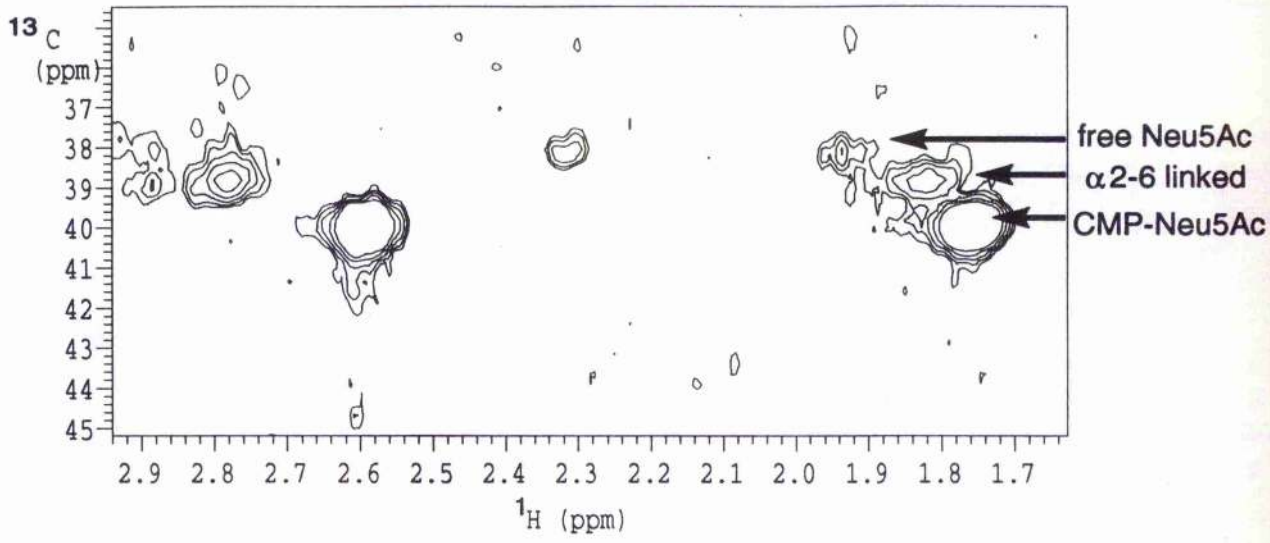


Figure 6.3: An extract from an ^{1}H-^{13}C HSQC spectrum of the crude resialylation reaction mixture, three pairs of cross-peaks can clearly be seen for Neu5Ac, CMP-Neu5Ac, and an α2,6 sialylated species.

A ^{13}C-^{1}H HSQC spectrum of purified [U-^{13}C] Neu5Acα2-6Gal labelled FcREA is shown in figure 6.4A. An estimated 80% of the glycan the possible acceptor sites were resialylated (based on the intensity of the residual asiaylated **6/6′** H6 intensities). It is clear from this spectrum that in common with the [U-^{13}C] galactosylated species, the Neu5Ac residues, **7** & **7′**, also experience two distinct chemical and mobile environments. Chemical shift differences between the 'bound' and 'free' shifts (assuming the pairs of peaks are a associated with **7′** and **7**, respectively) of the are of the order of 0.1-0.2 ppm, compared with the almost degenerate shifts (Δ 0.005 ppm) reported for the **6** and **6′**, and the **7** and **7′** resonances of the cleaved glycan (Schauer, R. 1982). To test the binding site hypothesis, the FcREA sample was heated to 65 °C, at this temperature there should be enough thermal energy to dissociate the glycan from the protein.

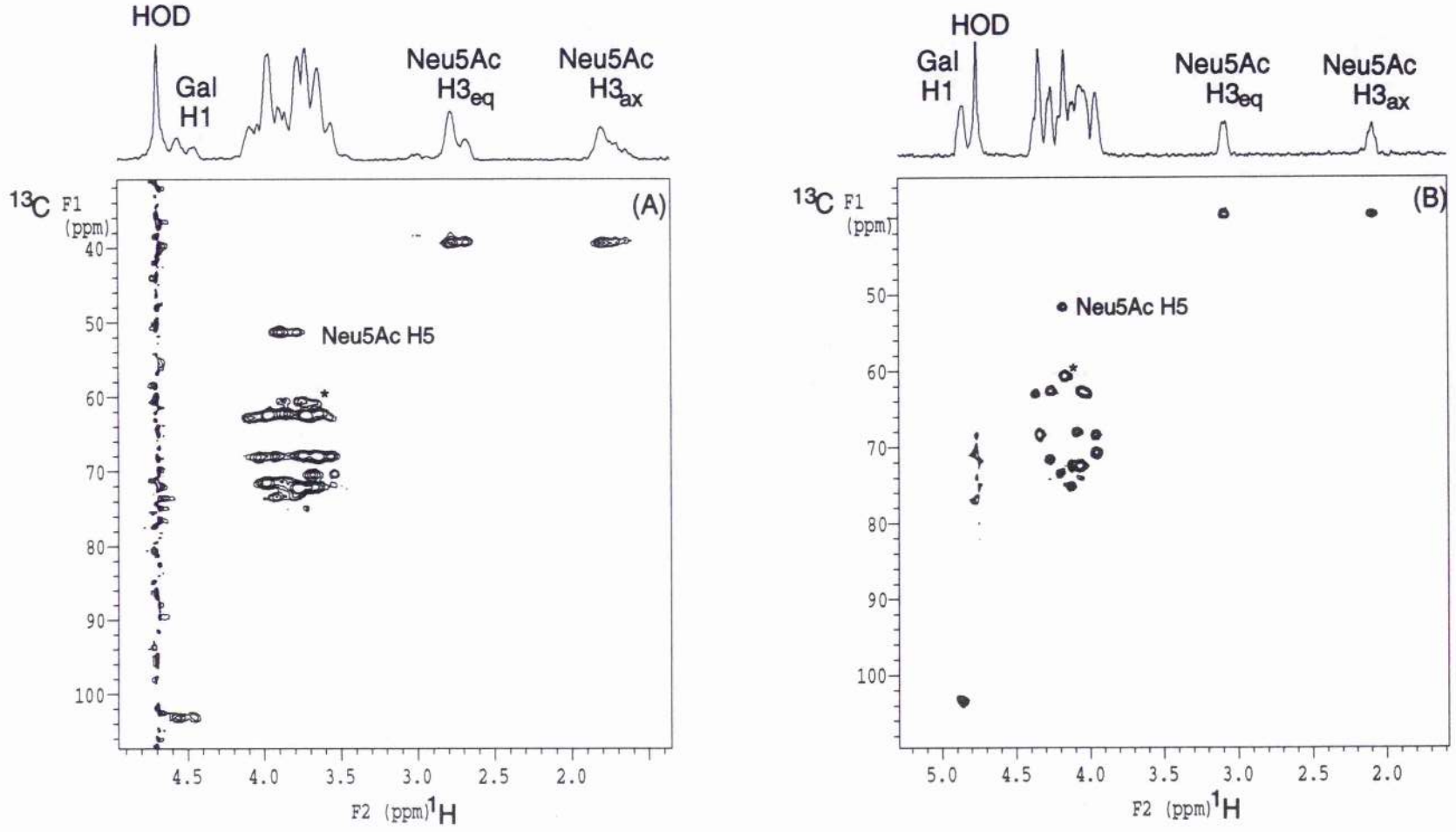


Figure 6.4: ^{13}C-^{1}H HSQC spectra of [U-^{13}C] Neu5Acα2-6Gal labelled FcREA at (A) 37 °C and (B) 65 °C.

A HSQC experiment acquired at this temperature (*See Figure 6.4B)* shows that the cross-peaks are no-longer split or significantly line broadened. The system was also completely reversible: cooling the sample back to 37 °C resulted in a spectrum (data not shown) identical to Figure 6.4A. These data are consistent with the model of an extended lectin patch for Neu5Ac.

Antibodies against the human Fc fragments have been isolated from rheumatoid arthritis patients. These RA associated factors may play a role in the auto-immune reaction known to cause swelling and tissue damage in the joints of RA patients. A preliminary X-ray crystal structure of a Fc - RA factor (F_{ab} KOK) has been solved (Lund, *personal communication, 1996*). However, the Fc glycan chains were not observed in the complex suggesting that the 6′ residue has been displaced from its 'binding' site and the resulting glycan is too mobile to give good electron densities. Heteronuclear NMR could be a powerful complementary tool as any change in the position of the 6′ residue would be evident in the NMR spectrum. Initial NMR experiments performed on the isolated F_{ab} KOK fragment revealed the presence of high concentrations of glucose (data not shown). The glucose signals could not be removed by exhaustive dialysis or repurification of the F_{ab} fragment. The F_{ab} fragment was isolated from a diabetic patient and it was concluded that the contaminating glucose signals were due to glycation, a phenomenon where proteins exposed to high concentrations of serum glucose are non-specifically glycosylated (Brownlee *et al.* 1988). NMR binding studies will provide essential information on the role of the glycan chain in Fc-RA factor complexes when alternative sources of RA factor are available.

6.4 Conclusions

It has been shown that a multienzymatic incorporation of *N*-acetyl [U-^{13}C] neuraminic acid to the terminal positions of an intact glycoprotein may be achieved in good yield (~80%) and without significant loss to proteolysis. Alternative methods for stable isotope labelling of intact glycoprotein glycans has been reported, such as expressing glycoproteins in CHO cells (Weller *et al.* 1996). However, in these expression systems the whole glycan chain and its associated protein are labelled and while it is possible to tailor NMR experiments to filter out protein resonances, significant leakage of protein signals has been reported and this approach may be unsatisfactory in some cases. In addition, the media required for such experiments is very expensive and as the protein may have to undergo several purification steps from the crude cell lysate to purified protein, where there is ample opportunity for the labile Neu5Ac glycosidic linkages to be cleaved. Therefore, the *relatively* less expensive method of selectively incorporating labelled sugars reported here, and the simplified spectra such samples would produce may be of more general use.

Finally, the ^{1}H-^{13}C HSQC spectrum of [U-^{13}C] Neu5Acα2-3Gal terminating FcREA suggest that both the **7′** Neu5Ac and **6′** galactose residues form a close associations with the C_H2 domain. Where such a binding site may exist on the surface of the C_H2 domain remains to be determined.

References

Acquotti, D. (1990). "3D structure of the oligosaccharide chain of G_{M1} ganglioside revealed by distance-mapping procedure: A rotating and laboratory frame NOE investigation of native glycolipid in dimethyl sulfoxide in water-dodecylphocholine solutions." *Journal of the American Chemical Society* **112** (21): 7772-7774.

Adams, B. and Lerner, L. (1992). "A simple one-dimensional method for measuring proton exchange rates in water." *Journal of Magnetic Resonance.* **96** (3) : 604-607.

Adler, Y., Lamour, A., Jamin, C., Menez, J. F., Lecorre, R., Shoenfeld, Y. and Youinou, P. (1995). "Impaired binding-capacity of asialyl and agalactosyl IgG to Fc-gamma receptors." *Clinical and Experimental Rheumatology* **13** (3): 315-319.

Amin, S. R. and Chatterjee, B. P. (1995). "Binding mechanism of methyl-α-*N*-acetyl D-galactopyranosamine to *Artocarpus lakoocha* lectin, artocarpin: A proton nuclear magnetic resonance study." *Draft Document* .

Andrews, J. S., Weimar, T., Frandsen, T. B., Svensson, B. and Pinto, B. M. (1995). "Novel disaccharides containing sulfur in the ring and nitrogen in the inter-glycosidic linkage - conformation of methyl 5'-thio-4-N-α-maltoside bound to glucoamylase and its activity as a competitive inhibitor." *Journal of the American Chemical Society* **117** : 10799-10804.

Arab, S., Khine, A. A., Rutka, S., Grinstein, S. and Lingwood, C. A. (1995). "Globotriosyl ceramide-mediated intracellular targeting of verotoxin. Retrograde transport to the nucleus and beyond." *Glycoconjugate Journal* **12** (4): 499.

Arepalli, S. R., Glaudemans, C. P. J., Daves, G. D., Kovac, P. and Bax, A. (1995). "Identification of protein-mediated indirect NOE effects in a disaccharide-Fab' complex by transferred ROESY." *Journal of Magnetic Resonance B* **106** : 195-198.

Asensio, J. L., Canada, F. J. and Jimenez-barbero, J. (1995). "Studies of the bound conformations of methyl α-lactoside and methyl β-allolactoside to ricin B chain using transferred NOE experiments in the laboratory frame and rotating frames, assisted by molecular mechanics and dynamics simulations." *European Journal of Biochemistry* **233** : 618-630.

Aitken, R. and Hirst, T. R. (1995). "Bacterial toxins as novel antigen delivery systems." *Livestock Production Science* **42** (2-3): 163-172.

Aubin, Y., Ito, Y., Paulson, J. C. and Prestegard, J. H. (1993). "Structure and dynamics of the sialic acid moiety of G_{M3}-ganglioside at the surface of a magnetically oriented membrane." *Biochemistry* **32** : 13045-13413.

Aubin, Y. and Prestegard, J. H. (1993). "Structure and dynamics of sialic acid at the surface of a magnetically oriented membrane system." *Biochemistry* **32** (13): 3422-3428.

Augé, C., David, S. and Gautheron, C. (1984). "Synthesis with immobilised enzyme of the most important sialic acid." *Tetrahedron Letters* **25** (41): 4663-4664.

Augé, C., Gautheron, C., David, S., Malleron, A., Cavaye, B. and Bouxom, B. (1990). "Sialyl aldolase in organic-synthesis - From the trout egg acid, 3-deoxy-D-glycero-D-galacto-2-nonulosonic acid (KDN), to branched-chain higher ketoses as possible new chirons." *Tetrahedron* **46** (1): 201-214.

Badet, B., Vermoote, P., Haumont, P.-Y., Lederer, F. and Goffic, F. L. (1987). "Glucosamine synthase from *Escherichia coli*: Purification, properties, and glutamine-utalising site location." *Biochemistry* **26** (7): 1940-1948.

Bains, G., Lee, R. T. and Freire, E. (1992). "Microcalorimetric study of wheat germ agglutinin binding to *N*-acetylglucosamine and its oligomers." *Biochemistry* **31** : 12624-12628.

Baisch, G., Ohrlein, R. and Ernst, B. (1996a). "Enzymatic glycosylation of non-natural glucosamide-acceptors." *Bioorganic & Medicinal Chemistry Letters* **6** (7): 749-754.

Baisch, G., Ohrlein, R., Streiff, M. and Ernst, B. (1996b). "Enzymatic α(2-3)-Sialylation of non-natural disaccharides with cloned sialyl-transferase." *Bioorganic and Medicinal Chemistry Letters* **6** (7): 755-758.

Baumberger, F. and Vasella, A. (1986). "Dexoy-nitosugars .15. Synthesis of

N-acetylneuraminic acid and *N*-acetyl-4-epinuraminic acid." *Helvetica Chimica Acta* **69** (5): 1205-1215.

Bax, A. (1994). "Multidimensional nuclear magnetic resonance methods for protein studies." *Current Opinion in Structural Biology* **4** : 738-744.

Bax, A., Clore, G. M., Dirscoll, P. C., Gronenborn, A. M., Ikura, M. and Kay, L. E. (1990). "Practical aspects of proton carbon - proton three-dimensional correlation spectroscopy of ^{13}C labelled proteins." *Journal of Magnetic Resonance* **87** : 620-627.

Bax, A., Delagio, F., Grzesiek, S. and Vuister, W. (1994). "Resonance assignment of methionine methyl groups and χ^3 angular information from long-range proton-carbon and carbon-carbon J correlation in a calmodulun-peptide complex." *Journal of Biomolecular NMR* **4** : 787-797.

Bax, A., Max, D. and Zax, D. (1992). "Measurement of long-range ^{13}C-^{13}C *J* couplings in a 20 kDa protein-peptide complex." *Journal of the American Chemical Society* **114** : 6923-6925.

Bednarski, M. D., Chenault, H. K., Simon, E. S. and Whitesides, G. M. (1987). "Membrane-enclosed enzymatic catalysis (MEEC): A useful, practical new method for the manipulation of enzymes in organic synthesis." *Journal of the American Chemical Society* **109** (4): 1283-1285.

Berg, J., Kroon-Batenburg, L. M. J., Strecker, G., Montreuil, J. and Vliegenthart, J. F. G. (1989). "Conformational analysis of the sialyl α(2-3/6)

N-acetyllactosamine structural element occurring in glycoproteins, by 2D NOE ^{1}H-NMR spectroscopy in combination with energy calculations by hard-sphere exo-anomeric and molecular mechanics force-field with hydrogen-bonding potential." *European Journal of Biochemistry* **178** : 727-739.

Berman, E., Brown, J. H., Lis, H. and Sharon, N. (1985). "Binding of [1-^{13}C] galactose-labelled N-acetyllactosamine to *Erythrina cristagalli* agglutinin as studied by ^{13}C-NMR." *European Journal of Biochemistry* **152** (2) : 447-451.

Bevilacqua, V. L., Kim, Y. and Prestegard, J. H. (1992). "Conformation of β–Methylmelibiose bound to the ricin B-chain as determined from transferred NOEs." *Biochemistry* **31** : 9339-9349.

Bevilacqua, V. L., Thomson, D. S. and Prestegard, J. H. (1990). "Conformation of methyl β-lactoside bound to the ricin B-chain: Interpretation of transferred NOE facilitated by spin simulation and selective deuteration." *Biochemistry* **29** (23): 5529-5537.

Bock, K., Arnarp, J. and Lonngren, J. (1982). "The preferred conformation of oligosaccharides derived from complex-type carbohydrate portions of glycoproteins." *European Journal of Biochemistry* **129** : 171-178.

Bock, K. and Lemieux, R. U. (1982). "The conformational properties of sucrose in aqueous solution - intramolecular hydrogen bonding."

Carbohydrate Research **100** : 63-74.

Bourne, Y., Rougé, P. and Cambillau, C. (1990a). "X-ray crystal structure of a Manα1-3Manβ1-4GlcNAc-lectin complex at 2.1 Å resolution." *Journal of Biological Chemistry* **265** : 18161-18165.

Bourne, Y., Rougé, P. and Cambillau, C. (1990b). "X-ray crystal structure of a biantennary octasaccharide-lectin complex at 2.3 Å resolution." *Journal of Biological Chemistry* **267** : 197-203.

Boyd, B., Tyrell, G., Maloney, M., Gyles, C., Brunton, J. and Lingwood, C. (1993). "Alteration of the glycolipid binding specificity of the pig edema toxin from globotetraosyl to globotriaosyl ceramide alters *in vivo* tissue targeting and results in a verotoxin 1-like disease in pigs." *Journal of Experimental Medicine* **177** (June): 1745-1753.

Boyd, P. N., Lines, A. C. and Patel, A. K. (1995). "The effect of the removal of sialic acid, galactose and total carbohydrate on the functionality of campath-1H." *Molecular Immunology* **32** (17-18): 1311-1318.

Brady, J. W. (1986). "Molecular dynamics simulations of α-D-glucose." *Journal of the American Chemical Society* **108** : 8153-8160.

Brady, J. W. (1987). "Molecular dynamics simulations of β-D-glucopyranose." *Carbohydrate Research* **165** : 306-312.

Brownlee, M., Cerami, A. and Vlassara, H. (1988) *New England Journal*

of Medicine : 1315-1321.

Bruschweiler, R., Griesinger, C., Sorensen, O. W., and Ernst, R. R. (1988). "Combined use of hard and soft pulses fore ω-1 decoupling in 2D NMR-spectroscopy." *Journal of Magnetic Resonance* **78** (1) : 178-185.

Bundle, D. R. (1992). "Carbohydrate-protein interactions in antibodies and lectins." *Current Opinion in Structural Biology* **2** : 666-673.

Bundle, D. R., Baumann, H., Brisson, J. R., Gagne, S. M., Zdanov, A. and Cygler, M. (1994). "Solution structure of a trisaccharide-antibody complex - comparison with a crystal structure." *Biochemistry* **33** : 5183-5192.

Burnette, N. W. (1994). "AB_5 ADP-ribosylating toxins: comparative anatomy and physiology." *Structure* **2** : 151-158.

Bystrov, V. F., Inanov, V. T., Portnova, S. L., Balashova, T. A. and Ovchinnikov, Y. A. (1973). "Refinement of the angular dependence of the peptide vicinal NH-$C^{\alpha}H$ coupling constant." *Tetrahedron* **29** : 873-877.

Campbell, A. P. and Sykes, B. D. (1993). "The two-dimensional transferred nuclear Overhauser effect: Theory and practice." *Annual Review of Biomolecular Structure* **22** : 99-122.

Carver, J. P., Michnick, A. I. and Cumming, D. A., Eds. (1989). Carbohydrate Recognition in Cellular Function. Ciba Foundation, Wiley-Interscience Publication.

Casset, F., Imberty, A., Perez, S., Etzler, M. E., Paulsen, H. and Peters, T. (1997). "Transferred nuclear Overhauser enhancement (NOE) and rotating frame NOE experiments on the Forssman pentasaccharide in the binding site of *Dolichos biflorus* lectin." *European Journal of Biochemistry* **244** : 242-250.

Cavanagh, J., Hunter, C. A., Jones, D. N. M., Keeler, J. and Saunders, J. K. M. (1988). "Practicalities and applications of reverse heteronuclear shift correlation - porphyrin and polysaccharide examples." *Magnetic Resonance in Chemistry* **26** : 867-875.

Clore, G. M. and Gronenborn, A. M. (1982). "Theory and applications of the transferred nuclear Overhauser effect to the study of the conformations of small ligands bound to proteins." *Journal of Magnetic Resonance* **48** : 402-417.

Clore, G. M. and Gronenborn, A. M. (1983a). "Theory of the time-dependant TRNOE: Applications to structural analysis of ligand-protein complexes in solution." *Journal of Magnetic Resonance* **53** : 443-442.

Clore, G. M. and Gronenborn, A. M. (1983b). "Theory of the time-dependent transferred nuclear Overhauser effect - application to analysis of ligand protein complexes." *Journal of Magnetic Resonance* **53** : 423-442.

Clore, G. M., Gronenborn, A. M., Brunger, A. T. and Karplus, M. (1985). "Solution conformation of a heptadecapeptide comprising the DNA binding helix-F of the cyclic-AMP receptor protein of *Escherichia coli* -

combined use of ^{1}H nuclear magnetic resonance and restrained molecular dynamics." *Journal of Molecular Biology* **186** : 435-455.

Cooke, R. M., Hale, R. S., Lister, S. G., Shah, G. and Weir, M. P. (1994). "The conformation of the sialyl Lewis X ligand changes upon binding to E-selectin." *Biochemistry* **33** : 10591-10596.

Cygler, M., Rose, D. R. and Bundle, D. R. (1991). "Recognition of a cell-surface oligosaccharide of pathogenic *Salmonella* by an antibody Fab fragment." *Science* **253** : 442-445.

Dabrowski, J., Kozar, T., Grosskurth, H., and Nifant'ev, N. E. (1995). "Conformational mobility of oligosaccharides: Experimental evidence for the existence of an 'anti' conformer of the Galβ1-3Glcβ-OMe disaccharide." *Journal of the American Chemical Society* **117** : 5534-5539.

de Graaf, F. K., Kelmm, P. and Gaastra, W. (1980). "Purification, characterisation and partial covalent structure of *Escherichia coli* adhesive antigen K99." *Infection and Immunology* **33** : 877-883.

Debye, P. (1929). Polar molecules. New York, Chemical Catalog Co.

DeGrandis, S., Law, H., Brunton, J., Gyles, C. and Lingwood, C. A. (1989). "Globotetraosylceramide is recognised by the pig edema disease toxin." *Journal of Biological Chemistry* **264** (21): 12520-12525.

Deisenhofer, J. (1981). "Crystallographic refinement and atomic models of

a human Fc fragment and its complex with fragment B of protein A from *Straphylococuus aureus* at 2.9 and 2.8Å resolution." *Biochemistry* **20** : 2361-2367.

Dutka-Malen, S., Mazodier, P. and Badet, B. (1988). "Molecular cloning and overexpression of the glucosamine synthase gene from *Eschericha coli.*" *Biochimie* **70** : 287-290.

Dweck, R. A., Lellouch, A. C. and Wormald, M. R. (1995). "Glycobiology: 'The function of sugar in the IgG molecule'." *Journal of Anatomy* **187** : 279-292.

Edelman, G. M., Cunningham, B. A., Gall, W. E., Gottlieb, P. D., Rutishauser, U. and Waxdal, M. J. (1969). "The covalent structure of an entire γG immunoglobulin molecule." *Proceedings of the National Academy of Science of the United States of America.* **63** : 78-85.

Edison, A. S., Abildgaard, F., Westler, W. M., Mooberry, E. S. and Markley, J. L. (1994). Practical introduction to theory and implementation of multinuclear, multidimensional nuclear magnetic resonance experiments. Methods in Enzymology. T. L. James and N. J. Oppenheimer. San Diego, Academic Press. INC. **239:** 3-79.

Einsphar, H. M. (1989). "Protein-carbohydrate interactions in biological systems." *Transactions of the American Crystallographic Association.* **25** : 1-22.

Farmer II, B. T., Macura, S. and Brown, L. R. (1987). "Relay artifacts in ROESY spectra." *Journal of Magnetic Resonance* **72** : 347-352.

Fishman, P. H. (1978). "Interaction of choleragen with the oligosaccharide portion of ganglioside G_{M1}: Evidence for multiple oligosaccharide binding sites." *Biochemistry* **17** (4): 711-716.

Fishman, P. H. (1980a). "Mechanism of action of cholera toxin: Effect of receptor density and multivalent binding on activation of adenylate cyclase." *Journal of Membrane Biology* **54** : 57-60.

Fishman, P. H. (1980b). "Mechanism of action of cholera toxin: Studies of the lag period." *Journal of Membrane Biology* **54** : 61-72.

Fishman, P. H. (1982). "Role of membrane gangliosides in the binding and action of bacterial toxins." *Journal of Membrane Biology* **69** : 85-97.

Fitz, W., Schwark, J.-R. and Wong, C.-H. (1995). "Aldotetroses and C(3)-modified aldohexoses as substrates for N-acetyl neuraminic acid aldolase: A model for the explanation of the normal and the inversed stereoselectivity." *Journal of Organic Chemistry* **60** : 3663-3670.

Forster, M. J. (1991). "Comparison of computational methods for simulating nuclear Overhauser effects in NMR spectroscopy." *Journal of Computational Chemistry* **12** : 292-300.

Fraser, R. R., Kaufman, M. and Morand, P. (1968). "Stereochemical dependence of vicinal H-C-O-H coupling constants." *Canadian Journal of*

Chemistry **47** : 403-409.

French, A. D. (1988). "Rigid and relaxed-residue conformational analyses of cellobiose using the computer program MM2." *Biopolymers* **27** : 1519-1525.

French, A. D. and Brady, J. W. (1989). Computer modelling of carbohydrate molecules. Washington., American Chemical Society.

Genest, D. (1989). "A Monte-Carlo simulation study of the influence of internal motions on the molecular conformation deduced from two-dimensional NMR experiments." *Biopolymers* **28** : 1903-1911.

Ghosh, S. and Roseman, S. (1962). "L-glutamine-D-fructose 6-phosphate transamidase from E.coli,." *Methods in Enzymology* **5** : 414-417.

Gilhespy-Muskett, A. M., Partridge, J., Jefferis, R. and Homans, S. W. (1994). "A novel 13C isotopic labelling strategy for probing the structure and dynamics of glycan chains in situ on glycoproteins." *Glycobiology* **4** (4): 485-489.

Goebel, C. V., Dimpfl, W. L. and Brant, D. A. (1970). "The conformational energy of maltose and amylose." *Macromolecules* **3** : 644-654.

Goodford, P. J. (1985). "A computational procedure for determining energetically favourable binding sites on biologically important macromolecules." *Journal of Medicinal Chemistry* **28** (7): 849-857.

Gordon, D. M. and Whitesides, G. M. (1993). "Indium-mediated allylations of unprotected carbohydrates in aqueous media: A short synthesis of sialic acid." *Journal of Organic Chemistry* **58** : 7937-7938.

Griffiths, S. L. (1986). "Characterisation of receptors for cholera toxin and E.coli heat-labile enterotoxin in rabbit intestinal brush borders." *Biochemical Journal* **238** : 313-322.

Grzesiek. S. and Bax, A. (1993). "Measurement of amide proton-exchange rates and NOEs with water in $^{13}C/^{15}N$-enriched calcineurin-B." *Journal of Biomolecular NMR.* **3** (6): 799-805.

Ha, S., Giammona, A. and Field, M. (1988a). "A revised potential-energy surface for molecular mechanics studies of carbohydrates." *Carbohydrate Research* **180** : 207-221.

Ha, S. N., Madsen, L. J. and Brady, J. W. (1988b). "Conformational analysis and molecular dynamics simulations of maltose." *Biopolymers* **27** : 1927-1952.

Haasnoot, C. A. G., de Leeuw, F. A. A. M. and Altona, C. (1980). "The relationship between proton-proton coupling constants and substituent electronegativities-1." *Tetrahedron* **36** : 2873-2792.

Hare, B. J., Rise, F., Aubin, Y. and Prestegard, J. H. (1994). "C-13 NMR-studies of wheat germ agglutinin interactions with N-acetylglucosamine at a magnetically oriented bilayer surface." *Biochemistry* **33** (33): 10137-

10148.

Harris, L. S., Larson, S. B., Hasel, K. W. and McPherson, A. (1997). "Refined structure of an intact IgG2a monoclonal antibody." *Biochemistry* **36** : 1581-1597.

Harris, R., Rutherford, T. J., Milton, M. J. and Homans, S. W. (1996). "Three-dimensional heteronuclear NMR techniques for assignment and conformational analysis using exchangeable protons in uniformly ^{13}C enriched oligosaccharides." *Biomolecular NMR* **9** : 47-54.

Harrison, J., Kartha, R., Field, R. A., Naismith, J. H. and Schenkman, S. (1997). "Development of a high through-put spectrophotometric assay to monitor*Trypanosoma cruzi trans*-sialidase." *Biochemical Society Transactions* **25** (48) 424S.

Heidlas, J. E., Watson, W. J., Pale, P. and Whitesides, G. M. (1992a). "Gram-scale synthesis of uridine 5'-diphosphoglucosamine: Comparison of enzymatic and chemical routes." *Journal of Organic Chemistry* **57** : 146-151.

Heidlas, J. E., Watson, W. J. and Whitesides, G. M. (1992b). "Pratical enzyme-based syntheses of uridine 5'-diphosphogalactose and uridine 5'-diphospho-*N*-acetylgalactosamine on a gram scale." *Journal of Organic Chemistry* **57** : 152-157.

Herrmann, G. F., Kragl, U. and Wandrey, C. (1993). "Continuous catalytic

synthesis of *N*-Acetyllactosamine." *Angewandte Chemie - International Edition in English* **32** (9): 1342-1343.

Hol, W. G. J., Sixma, T. K. and Merritt, E. A. (1995). "Structure and function of the *E. coli* heat-labile enterotoxin and cholera toxin B pentamer. Bacterial Toxins and Virulence Factors in Disease". J. Moss, B. Iglewski, M. Vaughan and A. T. Tu. New York, Marcel Dekker, Inc. **8:** 185-223.

Homans, S. W. (1990). "Oligosaccharide conformations: Application of NMR and energy calculations." *Prog. NMR. Spec.* **22** : 55-81.

Homans, S. W. (1993). "Conformation and dynamics of oligosaccharides in solution." *Glycobiology* **3** (6): 551-555.

Homans, S. W., Edge, C. J., Ferguson, M. A. J., Dwek, R. A. and Rademacher, T. W. (1989). "Solution structure of the Glycosylphosphatidylinositol membrane anchor glycan of *Trypanosoma bruciei* variant surface glycoprotein." *Biochemistry* **28** : 2881-2887.

Homans, S. W. and Forster, M. (1992). "Application of restrained minimisation, simulated annealing and molecular dynamics simulations of oligosaccharides." *Glycobiology* **2** (2): 143-151.

Homans, S. W., Pastore, A., Dwek, R. A. and Rademacher, T. W. (1987). "Structure and dynamics in oligo-mannose type oligosaccharides." *Biochemistry* **26** : 6649-6655

Howard, K. P. and Prestegard, J. H. (1995). "Membrane and solution conformations of monogalactosyldiacylglycerol using NMR/molecular modelling methods." *Journal of the American Chemical Society* **117** : 5031-5040.

Hwang, T.-L. and Shaka, A. J. (1995). "Water suppression that works. Excitation coupling using arbitrary waveforms and pulsed field gradients." *Journal of Magnetic Resonance, Series A* **112** : 275-279.

Ichikawa, Y. (1992). "Chemical-Enzymatic synthesis and conformational analysis of sialyl Lewis x and derivatives." *Journal of the American Chemical Society* **114** : 9283-9298.

Ichikawa, Y., Look, G. C. and Wong, C.-H. (1992). "Enzyme-catalyzed oligosaccharide synthesis." *Analytical Biochemistry* **202** : 215-238.

Ichikawa, Y., Shen, G.-J. and Wong, C.-H. (1991). "Enzyme-catalyzed synthesis of sialyl oligosaccharides with *in situ* regeneration of CMP-sialic acid." *Journal of the American Chemical Society* **113** (12): 4698-4700.

Imberty, A., Dalage, M.-M., Bourne, Y., Cambillau, C. and Perez, S. (1991). "Data bank of three-dimensional structures of disaccharides: Part II, N-acetyllactosamine type N-glycans." *Glycoconjugate Journal* **8** : 456-483.

Imberty, A., Tran, V. and Perez, S. (1989). "Relaxed potential energy surfaces of N-linked oligosaccharides: The mannose-α-1-3 mannose case." *Journal of Computational Chemistry.* **11** : 205.

Ito, I. and Paulson, J. C. (1993). "A novel strategy for the synthesis of ganglioside GM3 using an enzymatically produced sialoside glycosyl donor." *Journal of the American Chemical Society* **115** : 1603-1605.

Jacewicz, M., Clausen, H., Nudelamn, E., Donohue-Rolfe, A. and Keusch, G. (1986). *Journal of Experimental Medicine* **163** : 1391-1404.

Jackson, M. P. (1990). "Structure-function analysis of shiga toxin and the shiga-like toxins." *Microbial Pathogenesis* **8** : 235-242.

Janicot, M. (1991). "Activation of rat liver adenylate cyclase by cholera toxin internalisation and processing in endosomes." *Journal of Biological Chemistry* **266** (20): 12858-12865.

Karlsson, K.-A. (1995). "Microbial recognition of target-cell glycoconjugates." *Current Opinion in Structural Biology* **5** : 622-635.

Karlsson, K.-A., Angstrom, J., Bergstrom, J. and Lanne, B. (1992). "Microbial interaction with animal cell surface carbohydrates." *AMPIS, Supplement.* **100** (27): 71-83.

Karplus, M. (1959). "Contact electron-spin coupling of nuclear magnetic moments." *Journal of Chemical Physics* **30** (1): 11-18.

Karplus, M. (1963). "Vicinal proton coupling in nuclear magnetic resonance." *Journal of the the American Chemical Society.* **85** : 2870-2871.

Kittelmann, M., Klein, T., Kragl, U., Wandrey, C. and Ghisalba, O. (1992). "Preparative enzymatic synthesis of activated neuraminic acid by using a microbial enzyme." *Annals of the New York Academy of Sciences* **672** (November): 444-450.

Koketsu, M., Juneja, L. R., Kawanami, H., Kim, M. and Yamamoto, T. (1992). "Preparation of *N*-acetylneuraminic acid from delipidated egg-yolk." *Glycoconjugate Journal* **9** (2): 70-74.

Kong, D. C. M. and von Itzstein, M. (1995). "The first synthesis of a C-7 nitrogen containing sialic acid analogue, 5-acetoamido-7-azido-3,5,7-trideoxy-d-*glycero*-D-*galacto*-2-nonulopyranosonic acid (7-azido-7-deoxy-Neu5Ac)." *Tetrahedron Letters* **36** (6): 957-960.

Kragl, U. (1990). "Enzymatic synthesis of Sialic acid." *European Patent* **EP 0 428 947 A1** .

Kragl, U. (1992). "Reaktionstechnik bioatalystischer prozesse am beispiel der knotinuierlichen enzymatischen synthese von N-acetylneuraminsäure"., Institut für Biotechnologie.

Kragl, U., Appel, P., Gygax, D., Ghisalaba, O. and C., W. (1992). "Aldolases for use in carbohydrate synthesis: enzymatic reaction engineering as a tool for process optimisation". *Biochemical Engineering for 2001, Proceedings of Asia-Pacific Biochemical Engineering Conference*: 84-87.

Kragl, U., Gygax, D., Ghisalba, O. and Wandrey, C. (1991). "Enzymatic two-step synthesis of N-acetyl neuraminic acid in the enzyme membrane

reactor." *Angew.Chem.Int.Ed.Engl.* **30** (7): 827-828.

Kragl, U., Kittelmann, M., Ghisalaba, O. and C., W. (1995). "N-acetylneuraminic acid: from a rare chemical from natural sources to a multikilogram enzymatic synthesis for industrial application." *Annals of the New York Academy of Sciences (Enzyme engineering XII)* **750** : 300-305.

Kragl, U., Godde, A., Wandrey, C., Kinzy, W., Cappon, J. J. and Lutenburg, J. (1993). "Repetitive batch as an efficient method for preparative scale enzymatic synthesis of 5-azido-neuraminic acid and ^{15}N-L-glutamic acid." *Tetrahedron: Asymmetry* **4** (6): 1193-1202.

Kronis, K. A. and Carver, J. P. (1982). "Specificity of isolectins of wheat germ agglutinin for sialyloligosaccharides: A 360-MHz pmr binding study." *Biochemistry* **21** : 3050-3057.

Kronis, K. A. and Carver, J. P. (1985). "Wheat germ agglutinin dimers bind sialyloligosaccharides at four sites in solution: PMR temperature studies at 360MHz." *Biochemistry* **24** (826-833).

Lee, Y. C. (1972). "Analysis of sugars by automated liquid chromatography." *Methods in Enzymology* : 63-73.

Lemieux, R. U., Bock, K., Delbaere, L. T. J., koto, S. and Rao, V. S. (1980). "The human conformations of oligosaccharides related to the ABH human blood group determinants." *Canadian Journal of Chemistry.* **58** :

631-653.

Li, Y. C., and Montelione, G. T. (1993), "Solvent saturation-transfer effects in pulsed-field gradient heteronuclear single quantum coherence (PFG-HSQC) spectra of polypeptides and proteins." *Journal of Magnetic Resonance Series B* **101** (3): 315-325.

Lin-Chun Liu, J., Shen, G.-J., Ichikawa, Y., Rutan, J. F., Zapata, G., Vann, W. F. and Wong, C.-H. (1992). "Overproduction of CMP-Sialic acid synthase for organic synthesis." *Journal of the American Chemical Society* **114** (10): 3902-3910.

Lingwood, C. A. (1992). "Bacterial adhesions/Glycolipid receptors." *Current Opinion in Structural Biology* **2** : 693-700.

Lingwood, C. A. (1993). "Verotoxins and their glycolipid receptors." *Current Opinion in Structural Biology* **3**: 189-211.

Lis, H. and Sharon, N. (1991). "Lectin-carbohydrate interactions." *Current Opinion in Structural Biology* **1** : 741-749.

Livingston, P. O., Natoli, E. J., Jones-Calves, M., Stockbert, E., Oettgen, H. F. and Old, L. J. (1987). "Vaccines containing purified GM2 ganglioside elicit G_{M2} antibodies in melanoma patients." *Proceedings of the National Academy of Sciences of the United States of America.* **84** : 2911-2915.

Lommerse, J. P. M., Kroon-Batenberg, J. P. and Vliegenhart, J. F. G. (1995). "Conformational analysis of the of the xylose containing *N*-glycan of

pineapple stem bromelain as part of an intact glycoprotein." *Biochemsirty* **34** : 8196-8206.

London, R. E., Perlman, M. E. and Davis, D. G. (1992). "Relaxation-matrix analysis of the transferred nuclear Overhauser effect for finite exchange rates." *Journal of Magnetic Resonance* **97** : 79-98.

Low, D. G. (1996). NMR investigation of molecular contacts in a hapten-antibody complex. Chemistry. St. Andrews, University of St. Andrews: 210.

Lund, J. (1990). "A protein structural change in aglycosylated IgG3 correlates with loss of huFcyRi & huFcRiii binding and or/activation." *Molecular Immunology* **27** (11): 1145-1153.

Madsen, L. J., Ha, S. N., Trans, V. H. and Brady, J. W. (1990). Molecular dynamics simulations of carbohydrates and their solvation. Computer modelling of carbohydrate molecules.: 69-90.

Matsumoto, I., Koyama, T., Kitagaki-ogawa, H. and Seno, N. (1987). "Separation of isolectins by high-performance hydrophobic interaction chromatography." *Journal of Chromatography* **400** : 77-81.

McAuliffe, J. C. and Hindsgaul, O. (1997). "Carbohydrate drugs - an ongoing challenge." *Chemistry and Industry* **5** (3 March): 170-175.

Merrit, E. A. (1994). "Crytsal structure of cholera toxin B-pentamer bound to receptor GM1 pentasaccharide." *Protein Science* **3** : 166-175.

Milton, M. J. and Homans, S. W. (1995). "Thermodynamics and specificity of oligosaccharide interactions." *Glycoconjugate Journal* **12** (4): 429.

Milton, M. J., Harris, R., Probert, M. A., and Homans, S. W. (1997). "New structural restraints form carbon-13 enriched oligosaccharides" *Journal of the American Chemical Society*. Document submitted.

Morris, G. A. and Freeman, R. (1979). *Journal of the American Chemical Society* **101** : 760.

Mulloy, B., Frenkiel, T. A. and Davies, D. B. (1988). "Long-range carbon-proton coupling constants: Application to conformational studies of oligosaccharides." *Carbohydrate Research* **184** : 39-44

Murzin, A. G. (1993). "OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences." *EMBO* **13** (3): 861-867.

Naismith, J. H. and Field, R. A. (1996). "Structural basis of trimannoside recognition by concanavalin-A." *Journal of Biological Chemistry* **271** (2): 972-976.

Neuhaus, D. and Williamson, M. (1989). The nuclear Overhauser effect in structural and conformational analysis. London, VCH publishers.

Newkirk, M. M., Novick, J., Stevenson, M. M., Fournier, M. J. and Apostolakos, P. (1996). "Differential clearance of glycoforms of IgG in

hormal and autoimmune-prone mice." **106** (2): 259-264.

Ni, F. (1994). "Recent development in transferred NOE methods." *Progress in Nuclear Magnetic Resonance Spectroscopy* **26** (6): 517-606.

Ni, F. and Scheraga, H. A. (1994). "Use of transferred nOe to determine the conformations of ligands bound to proteins." *Accounts of Chemical Research* **27** (9): 257-264.

Nishida, Y., Hori, H., Ohrui, H. and Meguro, H. (1987). "^{1}H NMR analyses of rotameric distribution of C5-C6 bonds of D-glucopyranoses in solution." *Journal of Carbohydrate Chemistry* **7** : 239-250.

Noggle, J. H. and Schirmer, R. E. (1971). The nuclear Overhauser effect, chemical applications. New York, Academic Press.

Nyholm, P-G., Brunton, J. L. and Lingwood, C. A. (1995). "Modelling of the interaction of verotoxin-1 (VT1) with its glycolipid receptor, globotriosylceramide (Gb_3)." *International Journal of Biological Macromolecules* **17** (3-4): 199-204.

Nyholm, P-G., Magnusson, G. and Lingwood, C. A. (1996a). "Localization of two distinct globotriaosyl ceramide (Gb_3) binding sites in verotoxin1, 2, 2c and 2e by molecular modelling: Confirmation using deoxy analogues of Gb_3 and a new receptor for all verotoxins, de*N*-acetyl globotetraosyl ceramide (aminoGb_4)." *Document submitted* .

Nyholm, P-G., Magnusson, G., Zheng, Z., Norel, R., Binnington-Boyd, B. and Lingwood, C. A. (1996b). "Two distinct binding sites for globotriaosyl ceramide on verotoxins: identification by molecular modelling and confirmation using deoxy analogues and a new glycolipid receptor for all verotoxins." *Chemistry and Biology* **3** (April): 263-275.

O'Brien, A. D. and Holmes, R. K. (1987). "Shiga and shiga-like toxins." *Microbiological Reviews* **51** (2): 206-220.

Oettgen, H. F., Ed. (1989). Gangliosides and cancer. New York, VCH Publishers.

Olsen, S. and Phil, A. (1982). Molecular aspects of cellular regulation. Amsterdam, Elsevier Biochemical Press.

Otting, G. (1993). "Experimantal NMR techniques for studies of protein-ligand interactions." *Current Opinion in Structural Biology* **3** : 760-768.

Parekh, R. B. (1985). "Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG." *Nature* **316** .

Perkins, S. J. (1982). Application of ring current calculations to the proton NMR of proteins and transfer RNA. Biological Magnetic Resonance. L. Berliner and J. Reuben. New York, Plenum Press: 193-336.

Poppe, L., Dabrowski, J., von der Lieth, C.-W., Koike, K. and Ogawa, T. (1990). "Three-dimensional structure of the oligosaccharide terminus of

globotriaosylceramide and isoglobotriosylceramide in solution. A rotating-frame NOE study using hydroxyl groups as long-range sensors in conformational analysis by ^{1}H-NMR spectroscopy." *European Journal of Biochemistry* **189** : 313-324.

Poppe, L., Stuike-Prill, R., Meyer, B. and van Halbeek, H. (1992). "The solution conformation of sialyl-α(2-6)-lactose studied by modern NMR techniques and Monte Carlo simulations." *Journal of Biomolecular NMR* **2** : 109-136.

Poppe, L. and van Halbeek, H. (1991). "Nuclear magnetic resonance of hydroxyl and amino protons of oligosaccharides in aqueous solution: evidence for a strong intramolecular hydrogen bond in sialic acid residues." *Journal of the American Chemical Society* **113** : 363-365.

Poppe, L., von der Leith, C.-W. and J., D. (1990). "Conformation of the glycolipid globoside head group in various solvents in the micelle-bound state." *Journal of the American Chemical Society* **112** : 7762-7771.

Probert, M. A., Milton. M. J., Harris, R., Schenkman, S., Brown, J. M., Homans, S. W. and Field, R. A. (1997). "Chemoenzymatic synthesis of GM_3, Lewis x, and sialyl Lewis x oligosaccharides in ^{13}C-enriched form." *Tettrahedron Letters* in press.

Quiocho, F. A., Vyas, N. K. and Spurlino, J. C. (1989a). "Atomic interactions between proteins and carbohydrates." *Transactions of the American Crystallographic Association.* **25** : 23-35.

Quiocho, F. A., Wilson, D. K. and Vyas, N. K. (1989b). "Substrate specificity and affinity of a protein modulated by bound water molecules." *Nature* **340** : 404-407.

Rademaucher, T. W., Homans, S. W., Parekh, R. B. and Dwek, R. A. (1985). "Immunoglobulin G as a glycoprotein." *Biochemistry Society Symposium* **51** : 131-148.

Reuter, G. and Gabius, H.-J. (1996). "Sialic acids. Structure-analysis-metabolism-occurence-recognition." *Biological Chemistry* **377** (June): 325-342.

Richardson, J. M., Evans, P. D., Homans, S. W. and DonohueRolfe, A. (1997). "Solution structure of the carbohydrate-binding B-subunit of verotoxin VT-1 from E.coli." *Nature Structural Biology* **4** (3): 190-193.

Richardson, J. M., Milton, M. J. and Homans, S. W. (1995). "Solution dynamics of the oligosaccharide moiety of ganglioside G_{M1}: Comparison of solution conformations with the bound state conformation in association with cholera toxin B-pentamer." *Journal of Molecular Recognition* **8** : 358-362.

Roden, L., Roden, M., Yu, H., Jin, J. and Greenshields, J. (1993). "Separation of sugars by ion-exclusion chromatography on a cation resin." *Journal of Chromatography* **638** (1): 29-34.

Rougvie, R. (1997). Breakthrough in drive to stop deaths from *E. coli.* The

Scotsman. Edinburgh: 10.

Rutenber, E. and Robertus, J. D. (1989). "Carbohydrate interactions with ricin." *Transactions of the American Crystallographic Association.* **25** : 77-85.

Rutherford, T. J. and Homans, S. W. (1994). "Restrained vs. free dynamics simulations of oligosaccharides: Application to solution dynamics of biantennary and bisected biantennary N-linked glycans." *Biochemistry* **33** : 9606-9614.

Rutherford, T. J., Spackman, D. G., Simpson, P. J. and Homans, S. W. (1994). "5 Nanosecond molecular dynamics and NMR study of conformational transitions in the sialyl-Lewis X antigen." *Glycobiology* **4** (1): 59-68.

Sabesan, S. (1984). "The conformational properties of the gangliosides G_{M2} and G_{M1} based on ^{1}H ^{13}C NMR studies." *Canadian Journal of Chemistry* **62** : 1034-1045.

Sabesan, S. and Paulson, J. C. (1986). "Combined chemical and enzymatic synthesis of sialyloligosaccharides and characterisation by 500-MHz ^{1}H and ^{13}C NMR spectroscopy." *Journal of the American Chemical Society* **108** (8): 2068-2080.

Sanchez, J., Argottte, R. and Buela, A. (1997). "Engineering of cholera toxin A-subunit for carriage of epitopes at its amino end." *FEBS Letters*

401 (1): 665-678.

Sanders, J. K., and Hunter, B. K. (1990). "Modern NMR Spectroscopy - A Guide for Chemists." Oxford University Press, Oxford.

Sanders, C. R. and Prestegard, J. H. (1991). "Orientation and dynamics of β-dodecyl glucopyranoside in phospholipid bilayers by orientated sample NMR and order matrix analysis." *Journal of the American Chemical Society* **113** (6): 1987-1996.

Sanders, C. R. and Prestegard, J. H. (1992). "Headgroup orientations of alkyl glycosides at a lipid bilayer interface." *Journal of the American Chemical Society* **114** (18): 7096-7107.

Sastry, M. V. K., Swamy, M. J. and A., S. (1988). "Analysis of dynamics and mechanism of ligand binding to *Artocarpus integrifolia* agglutinin. A ^{13}C and ^{19}F NMR study." *Journal of Biological Chemistry* **263** (29): 14826-14381.

Saunders, J. K. M. and Hunter, B. K. (1987). Modern NMR spectroscopy: A guide for chemists. Oxford, Oxford University Press.

Scarsdale, J. (1988). "A molecular mechanics-NMR Pseudoenergy Approach to the solution conformation of glycolipids." *Journal of Computational Chemistry* **9** (2): 134-147.

Scarsdale, J. N., Yu, R. K. and Prestegard, J. H. (1986). "Structural analysis

of a glycolipid head group with one and two-state NMR pseudoenergy approaches." *Journal of the American Chemical Society* **108** : 6778-6784.

Schauer, R. (1982). Sialic acids-Chemistry, metabolism and function. Wein, Springer-Verlag.

Scheffler, K., Ernst, B., Katapodis, A., Magnani, J. L., Wang, W. T., Weiseman, R. and Peters, T. (1995). "Determination of the bioactive conformation of the carbohydrate ligand in the E-selectin in sialyl Lewis ^{x} complex." *Angewandte Chemie - International Edition in English* **34** : 1841-1844.

Schengrund, C. L. and Ringler, N. J. (1989). "Binding of *Vibro cholera* toxin and heat labile enterotoxin of *E.coli* to G_{M1} derivatives of G_{M1} and non-lipid polyvalent ligands." *Journal of Biological Chemistry* **264** (22): 13233-13237.

Scudder, P., Doom, J. P., Chuenkova, M., Manger, I. D. and Pereira, M. E. A. (1993). "Enzymatic characterisation of β–D-galactoside α2,3–*trans*-sialidase from *Trypanosoma cruzi*." *Journal of Biological Chemistry* **13** (5): 9886-9891.

Sharon, N. and Halina, L. (1990). "Carbohydrate-protein interactions." *Chemistry in Britain* : 679-682.

Sharon, N. and Lis, H. (1993). "Carbohydrates in cell recognition." *Scientific American* **268** (1): 74-81.

Siebert, H-C, Reuter, G., Schauer, R., von der Lieth, C.-W. and Dabrowski, J. (1992). "Solution conformations of G_{M3} gangliosides containing different sialic acid residues as revealed by NOE-based distance mapping molecular mechanics, and molecular dynamics calculations." *Biochemistry* **31** : 6962-6971.

Simon, E. S., Bednarski, M. D. and Whitesides, G. (1988). "Synthesis of CMP-NeuAc from N-Acetyl glucosamine: Generation of CTP from CMP Using Adenylate Kinase." *Journal of the American Chemical Society* **110** : 7159-7163.

Sixma, T. K., Pronk, S. E., Kalk, K. H., Wartna, E. S., van Zanten, B. A. N., Witholt, B. and Hol, W. G. (1992). "Lactose binding to heat-labile enterotoxin revealed by X-ray crystallography." *Nature* **355** : 561-564.

Sixma, T. K., Stein, P. E., Hol, W. G. J. and Read, R. J. (1993). "Comparison of the B-pentamers of heat-labile enterotoxin and verotoxin-1: two structures with remarkable similarity and dissimilarity." *Biochemistry* **32** : 191-198.

Spangler, B. D. (1992). "Structure and function of cholera toxin and the related *E.coli* heat-labile enterotoxin." *Microbiology Reviews* **56** (4): 622-648.

Sparks, M. A., Williams, K. W., Lukacs, C., Schrell, A., Priebe, G., Spaltenstein, A. and Whitesides, G. M. (1993). "Synthesis of potential inhibitors of hemagglutination by influenza virus: Chemoenzymatic

preparation of N-5 analogues of *N-acetyl neuraminic* acid." *Tetrahedron* **49** (1): 1-12.

Spivak, C. T. and Roseman, S. (1959). "Preparation of N-acetyl-D-mannosamine and D-mannosamine hydrochloride." *Journal of the American Chemical Society* **81** : 2403-2404.

Spurlino, J. C., Rodseth, L. E. and Quiocho, F. A. (1992). "Atomic interactions in Protein-Carbohydrate Complexes Tryptophan residues in the periplasmic maltodextrin receptor for active transport and chemotaxis." *Journal of Molecular Biology* **226** : 15-22.

States, D. J. Haberkon, R. A. and Ruben, D. J. (1982) "A two-dimensional nuclear Overhauser experiments with pure absorption phase in four quadrants." Journal of Magnetic Resonance **48** :, 2: 286-292

Stein, P. E., Boodhoo, A., Tyrrell, J. G., Brunton, J. L. and Read, R. J. (1992). "Crystal structure of the cell-binding B oligomer verotoxin-1 from *E.coli.*" *Nature* **355** : 748-749.

Stoddard, B. L. and Koshland Jnr, D. E. (1992). "Prediction of the structure of a receptor-protein complex using a binary docking method." *Nature* **358** : 774-776.

Subramanian, S. and Bax, A. (1987). "Generation of pure phase NMR subspectra for measurement of homonuclear coupling-constants." *Journal of Magnetic Resonance* **71** (2): 325-330.

Swift, T. J. and Connick, R. E. (1962). "NMR-relaxation mechanisms of ^{17}O in aqueous solutions of paramagnetic cations and lifetime of water molecules in the first co-ordination sphere." *Journal of Chemical Physics* **37** (2): 307-320.

Takahashi, N. and Ikuko, I. (1987). "Comparative structural study of the N-linked oligosaccharide moieties of human normal and pathological IgG." *Biochemistry* **26** : 1137-1144.

Teneberg, S., Hirst, T. R., Angstrom, J. and Karlsson, K.-A. (1994). "Comparison of the glycolipid-binding specificities of cholera toxin and porcine *Escherichia coli* heat-labile toxin: identification of a receptor-active non-ganglioside glycolipid for the heat-labile toxin in infant rabbit small intestine." *Glycoconjugate Journal* **11** : 533-540.

Thogerson, H., Lemieux, R. U., Bock, K. and Meyer, B. (1982). "Further justification for the exo-anomeric effect. Conformational analysis based on nuclear magnetic resonance spectroscopy of oligosaccharides." *Canadian Journal of Chemistry* **60** : 44-57.

Torda, A. E., Ruud, R. M. and van Gunsteren, W. F. (1990). "Time-averaged NOE distance restraints applied to tendamistat." *Journal of Molecular Biology* **214** : 223-235.

Tropp, J. (1980). "Dipolar relaxation and nuclear Overhauser effects in nonrigid molecules: The effect of fluctuating internuclear distances." *Journal of Chemical Physics* **72** : 6035-6044.

Tvaroska, I., Hricovini, M. and Petrakova, E. (1989). "An attempt to derive a new Karplus-type equation of vicinal proton-carbon coupling constants for C-O-C-H segments of bonded atoms." *Carbohydrate Research* **189** : 359-362.

Tvaroska, I. and Bleha, T. (1989). "Anomeric and exo-anomeric effects in carbohydrate chemistry. Advanced in carbohydrate chemistry and biochemistry." *Biochemistry*. **47** : 45-123.

Tyrrel, G. J. (1992). "Alteration of carbohydrate binding specificity of verotoxins from Galα1-4 to GalNacβ1-3Galα1-4Gal and vice versa by site-directed mutagenesis of the binding subunit." *Proceedings of the National Academy of Science of the United States of America.* **89** : 524-528.

Umenmoto, K., Oikawa, S., Aida, M. and Sugawara, Y. (1988). "Intermolecular NOE and atomic pair potential approaches to wheat germ agglutinin-sugar binding." *Journal of Biomolecular Structure & Dynamics* **6** (3): 593-608.

van Halbeek, H. (1996). Carbohydrates & glycoconjugates. Encyclopaedia of NMR spectroscopy: 1133-1137.

van Heynigen, W. E. (1971). "Deactivation of cholera toxin by ganglioside." *Journal of Infectious Disease* **124** : 415-418.

Vark, A. (1993). "Biological roles of oligosaccharides: all of the theories are correct." *Glycoconjugates* **3** (2): 98-130.

Varki, A. (1997). "Sialic acids as ligands in recognition phenomena." *FASEB Journal* **11** : 248-255.

von Itzstein, M., Wu, W.-Y. and Kok, G. B. (1993). "Rational design of potent sialidase-based inhibitors of influenza virus replication." *Nature* **363** : 418-423.

Vyas, N. K. (1991). "Atomic features of protein-carbohydrate interactions." *Current Opinion in Structural Biology* **1** : 732-740.

Watzele, G. and Tanner, W. (1989). "Colning of the glutamine: fructose-6-phosphate amidotransferase gene from yeast. Pheromonal regulation of its transcription." *Journal of Biological Chemistry* **264** (15): 8753-8758.

Weimer, T. and Peters, T. (1994). "*Aleuria-aurantia* agglutinin recognises multiple conformations of α-L-Fuc-(1-6)-β-D-GlcNAx-ome." *Angewandte Chemie - International Edition in English* **33** : 88-91.

Weller, C. T. (1994). Studies on the conformation and dynamics of the lipophosphoglycan of *Leishmania major.* Biochemistry. Dundee, University of Dundee: 226.

Weller, C. T., Lustbader, J., Seshadri, K., Brown, J. M., Chadwick, C. A., Kolthoff, C. E., Ramnarain, S., Pollak, S., Canfield, R. and Homans, S. W.

(1996). "Structural and conformational analysis of glycan moieties *in situ* on isotopically ^{13}C, ^{15}N-enriched recombinant Human chorionic gonadotropin." *Biochemistry* **35** (27): 8815-8823.

Wright, C. S. (1980). "Crystallographic elucidation of the saccharide binding mode in wheat germ agglutinin and its biological significance." *Journal of Molecular Biology* **141** : 267-291.

Wright, C. S. (1984). "Structural comparison of the two distinct sugar binding sites in wheat germ agglutinin isolectin 2." *Journal of Molecular Biology* **178** : 501-529.

Wright, C. S. (1987). "Refinement of the crystal structure of wheat germ agglutinin isolectin 2 at 1.8Å resolution." *Journal of Molecular Biology* **194** : 501-529.

Wright, C. S. (1990). "2.2 Å resolution structure analysis of two refined *N*-acetylneuraminyl-lactose wheat germ agglutinin isolectin complexes." *Journal of Molecular Biology* **215** : 635-651.

Wright, C. S. (1992). "Crystal structure of a wheat germ agglutinin/glycophorin-sialoglycopeptide receptor complex. Structural basis for cooperative lectin-cell binding." *Journal of Biological Chemistry* **267** (15): 14345-14352.

Wright, C. S. and Jäger, J. (1993). "Crystallographic refinement and structure analysis of the complex of wheat germ agglutinin with a

bivalent sialoglyco-peptide from glycophorin A." *Journal of Molecular Biology* **232** : 620-638.

Wright, C. S. and Kellog, G. E. (1996). "Differences in hydropathic properties of ligand binding at four independent sites in wheat germ agglutinin-oligosaccharide crystal complexes." *Protein Science* **5** : 1466-1476.

Wyss, D. F., Choi, J. S. and Wagner, G. (1995) Composistion and sequence specific resonance assignment of the heterogeneous *N*-linked glycan in the 13.6 kDa adhesive domain of human CD2 determined by NMR of the intact glycoprotein." *Biochemsitry* **34** : 1622-1634.

Yamamoto, T., Teshima, T., Inami, K. and Shiba, T. "Synthesis of sialic acid through aldol condensation of glucose with oxalacetic acid." *Carbohydrate Research* : 325-327.

Yan, Z.-Y. and Bush, C. A. (1990). "Molecular dynamics simulations and the conformational mobility of blood group oligosaccharides." *Biopolymers* **29** : 799.

Yu, L., Goldman, R., Sullivan, P., Walker, G. F. and Feisk, S. W. (1993). "Heteronuclear NMR of ^{13}C-labelled yeast cell wall β–glucan oligosaccharides." *Journal of Biomolecular NMR* **3** : 429-441.